Brain and Language **75,** 34–46 (2000)
doi:10.1006/brln.2000.2323, available online at http://www.idealibrary.com on IDEAL®

Pitch and Timing Abilities in Inherited Speech and Language Impairment

Katherine J. Alcock, Richard E. Passingham, Kate Watkins*, and Faraneh Vargha-Khadem*

*Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom; *Cognitive Neuroscience Unit, Institute of Child Health, University College London Medical School, and Great Ormond Street Hospital for Children, London, United Kingdom*

Members of the KE family who suffer from an inherited developmental speech-and-language disorder and normal, age-matched, controls were tested on musical abilities, including perception and production of pitch and rhythm. Affected family members were not deficient in either the perception or production of pitch, whether this involved either single notes or familiar melodies. However, they were deficient in both the perception and production of rhythm in both vocal and manual modalities. It is concluded that intonation abilities are not impaired in the affected family members, whereas their timing abilities are impaired. Neither their linguistic nor oral praxic deficits can be at the root of their impairment in timing; rather, the reverse may be true. © 2000 Academic Press

INTRODUCTION

Members of the KE family have a developmental disorder of speech and language that is inherited as an autosomal dominant trait (Fisher et al., 1998). Fifteen members of the extended family of 30 are affected and have both expressive and receptive speech and language difficulties. They also have difficulties with nonverbal oral movements (Vargha-Khadem et al., 1995; Alcock et al., submitted), resembling those found in adult patients with dysphasia resulting from a left-hemisphere cerebrovascular accident (Alcock, 1995; Alcock et al., submitted). The current study examines additional functions that may be impaired in this inherited speech and language disorder, namely pitch and timing abilities.

Evidence has accumulated recently indicating that fine-grained temporal processing is fundamental to speech and language. Most such investigations

Address correspondence and reprint requests to Katherine J. Alcock, Psychology Department, City University, Northampton Square, London EC1V 0HB. E-mail: k.j.alcock@city.ac.uk.

0093-934X/00 $35.00

have been carried out in the area of perception. Thus, in many studies of adults with acquired left-hemisphere brain damage (Hammond et al., 1982; Peretz, 1990; Robin et al., 1990; Phillips & Farmer, 1990), as well as of subjects with developmental dyslexia or specific language impairment (SLI) (Tallal et al., 1989, 1991), the affected individuals were found to be impaired in the discrimination of fine-grained temporal stimuli involving both music and speech. This evidence has been linked to differences that have been observed in normal subjects between the processing capabilities of the left and right hemispheres for stimuli with subtle temporal differences. Indeed, Tzeng and Wang (1984) have suggested that the left hemisphere's dominance for both speech and language and fine motor control stems from its specialization for fine-grained temporal processing.

Other studies have begun to examine the neuropsychology of timing in the production of movements (O'Boyle et al., 1990; Penhune et al., 1998). For example, Halsband et al. (1993) showed that lesions in the left supplementary motor cortex lead to an impairment in the reproduction of manual rhythms. One role of the motor cortex serving speech may be analogous, to coordinate oral movements with the temporal resolution needed for normal speech production.

The perception and production of pitch, or fundamental frequency, is another important component not only of musical ability but also of language, where it is part of the stress system of words and sentences. In English, as well as in tonal languages such as Chinese, it can be used to indicate semantics and syntax. Some individuals who have suffered right-hemisphere damage show impairment in pitch perception (Shankweiler, 1966; Zatorre, 1985; Samson & Zatorre, 1991; Peretz, 1990) or production (Damasio et al., 1975; Botez & Wertheim, 1959; Speedie et al., 1993), and this impairment could be related to the abnormal prosody that has been seen in adults with right hemisphere lesions.

The association of impairment in certain aspects of speech and language with an impairment in timing and perhaps also in intonation abilities suggests the possibility that the affected members of the KE family may likewise be impaired in timing and intonation abilities. To investigate this possibility, we tested the affected members and controls on intonation (both pitch and familiar melodies) and rhythm (complex temporal patterns) using measures of both musical perception and production.

METHODS

Subjects

Sixty subjects participated, 9 affected family members (mean age = 28.7, *SD* = 22.9) and 51 normal controls (mean age = 18.3, *SD* = 11.4). The control group, composed of different subgroups matched in age to different generations of the affected family members, included children attending one of four different schools in three varied areas of Oxford, adult volunteers

TABLE 1
Neuropsychological Data from Affected Family Members and Controls

Subject	Sex	Age	Years of education completed	Musical experience score	MLU on bus story	Score on animal naming task
AF3	M	14		3	6.2	23
AF4	M	11		3		15
AF5	F	11		3		25
AF6	F	10		3		16
AF7	F	41	11	3	8.6	33
AF8	M	21	11	3	7.1	21
AF9	F	48	11	3	7.3	26
AF10	F	78	9	3	6.6	14
AF11	M	44	11		7.2	14
M		28.67	10.6	3.00	7.17	20.25
SD		22.92	0.89	0.00	0.82	6.84
Controls, *M*		18.2716	11.35	3.43		
Controls, *SD*		11.3941	1.07	0.99		

from the Department of Experimental Psychology's Subject Panel, and adults recruited through an employment agency. Data on screening tests for all family members together with summary data for the control subjects are shown in Table 1.

Procedures

Brief descriptions of the procedures for each task and of the instructions to the subjects are given below. Full descriptions are presented in Alcock (1995).

Perception Tasks

For the music-perception tasks, tapes were prepared using a music-processing package (MusicPrinterPlus and Notator) to ensure that stimuli were consistent across subjects. There were three discrimination tasks.

Pitch discrimination. Two musical notes were presented in succession, with half the pairs being the same and with the frequency difference in the other half ranging from one semitone to a major 7th. This test had 22 items.

Melody discrimination. Familiar melodies were selected on the basis of a questionnaire given to subjects of the appropriate age groups to determine which songs were best known. Each melody was presented, followed by either an identical rendering or an alteration. Alterations were either a contour violated version (Peretz, 1990) or a novel version in which only the rhythm and the first two and last two notes were retained.

Rhythm discrimination. Rhythms were presented in pairs, half of which did not differ and in half of which in the second rhythm the length of notes was altered without changing the number of notes or the overall time signature (Peretz, 1990).

The tests were administered by the experimenter (K.J.A.) using a set-up consisting of two pairs of headphones and one tape player, which allowed the experimenter to hear what the subject heard, adjust volume levels, pause when appropriate, and prompt the subject for a response when none was forthcoming. On each trial, subjects indicated their choice nonverbally by pointing to one of two cards labeled SAME and DIFFERENT. On each task, the first two trials were presented as practice trials, with the subject given feedback after each response as to whether it was correct and, if not, what it should have been. After these first two trials, the only feedback given was of the most general kind (e.g., ‘‘you’re doing fine’’ and ‘‘yes, you seem to have the right idea’’). Further details of the perception tasks, including musical scores of the items used, can be found in Alcock (1995).

Production Tasks

Subjects were tested on the production tasks in a soundproof room at Great Ormond Street Hospital for Children, London, or in the Department of Experimental Psychology, Oxford, or, for some of the control children, in a quiet room at their school. Vocal output was recorded using an Altai UD130 microphone and a Marantz CP230 tape recorder. Manual output was recorded with a flat microphone (Realistic PS). Scoring was based on analysis of the tapes, as described below. There were three tests of production.

Pitch production. The subject was asked to sing individual notes after the experimenter had sung them. The experimenter started on a note in a comfortable range and sang 10–15 notes within the major scale having the initial note as its tonic, but choosing higher or lower notes as seemed most appropriate for the subject’s vocal range. The experimenter encouraged the subject to repeat each one, giving an appropriate cue if the subject anticipated or failed to sing, but providing no feedback as to accuracy.

Melody production. For each age group (under 30 and 30 or over), a different set of 10 songs with words and five melodies without words (or without well-known words) were selected (see Alcock, 1995, for a list). Subjects were first asked to sing as much as they knew of each of the songs with words, given the title and/or first line. It was emphasised that because the most important part of the song was the tune, the subject should sing the tune without words rather than stop if they knew no more words. After the list of songs with words was tested once, any songs that a subject had failed to sing were cued by the experimenter singing the first line or, if necessary, two lines. Subjects were again encouraged to sing the tune alone if the words were not known. The melodies without words (or without well-known words) were tested next, with the experimenter singing the first line or more if necessary. Subjects who knew words were encouraged to sing them to assist with recall.

Rhythm production. There were 14 rhythms, each consisting of four to nine notes. These were shortened versions of selected rhythms taken from the discrimination task (shown in Alcock, 1995). The first part of the task consisted of the experimenter tapping out each rhythm on the table and then asking the subject to manually reproduce each one immediately after presentation. The second part of the test consisted of the experimenter vocalizing the same rhythms to the phoneme /p/ and asking the subject to repeat each one vocally.

Scoring

The responses of 10% of the subjects were scored independently by two raters. Mean rate of agreement was 84.2%. In addition, for the production tasks, the tape recordings of the subjects’ single notes and of both their manual and vocal rhythms were digitized using the Waves+ program, part of the ESPS system, running on a Sun Workstation in the Oxford University Phonetics Laboratory. The digitized sound samples were then processed using one of two other ESPS programs, formant or get_f0, both of which extract fundamental frequency. Pitch trace files were converted into ASCII data files, and parameters including the mean pitch, range of pitch, and a number of pitch transitions were calculated for each sample. For single notes, the following parameters were calculated: mean, minimum and F0 in Hz, and the number of semitones by which the reproduced note differed from the note given as a model. For rhythms, the length of the interval between one note and the next was calculated, and these were then expressed as a proportion of the shortest interval produced. The intervals were then recalculated so that the unitary interval was the mean of all those intervals rounded down to a proportion of 1 rather than the shortest interval overall. The pattern of intervals produced on each item was then compared with the pattern given as a model and scored for the numbers of correct notes and numbers of correct rhythms.

Each pitch reproduction was scored as correct if the note differed by less than one semitone from the model provided by the experimenter; for incorrect items, the size of the error in semitones was calculated.

Each melody reproduction was rated, first, according to whether the subject had (a) attempted the song in response to the title and/or first line (not applicable to the melodies without words), (b) attempted the song or melody with assistance in the form of a cue, or (c) failed to sing the item. Second, each response was scored for the total number of notes sung as well as for the number of notes needed to cue the response. Finally, the first 10 notes of each song were scored in detail as follows: The pitch transition between each pair of notes was scored as correct if it matched the modal pitch transition for all subjects on that pair; if it did not match, the error was calculated as the difference in semitones between the actual pitch transition and the modal pitch transition.

Each rhythm reproduction was scored for whether it was an exact copy of the model and, if not, for how many notes were reproduced correctly, regardless of the position of any missing notes. For example, if the second of five notes was missing, the subject was scored as producing four notes correctly of five rather than producing only the first note correctly in the correct position.

RESULTS

Perception Tasks

Pitch and melody discrimination (Figs. 1 and 2). The two groups did not differ on either of these tasks.

Rhythm discrimination (Fig. 2). The affected family members discriminated fewer rhythms than the control group (two-tailed test: $t = 1.98$, $df = 56$, $p = .052$; one-tailed test: $p < .05$).

Production tasks

Pitch production (Fig. 3). The two groups did not differ.

Melody production (Figs. 4 and 5). The two groups did not differ either on the songs or on the melodies without words with respect to any of these measures: level of success (number of songs sung spontaneously vs number sung with cueing), number of notes sung, number of notes needed as a cue, or number of semitones by which the actual pitch transitions between note pairs differed from the modal pitch transitions.

Rhythm production (Figs. 6 and 7). Based on the rating scores, the affected family members reproduced fewer rhythms than the controls both manually ($F = 53.29$, $df = 1$, $p < .001$) and vocally ($F = 28.53$, $df = 1$, $p < .001$). MANOVAs carried out on the digitized data for number of rhythms reproduced completely revealed, in addition, that both groups performed better on manual than on vocal rhythms (Modality: $F = 10.38$, $df = 1$, $p = .005$) and that the affected family members were poorer than the controls on the manual rhythms only (Group by Modality: $F = 4.12$, $df = 1$, $p = .05$).

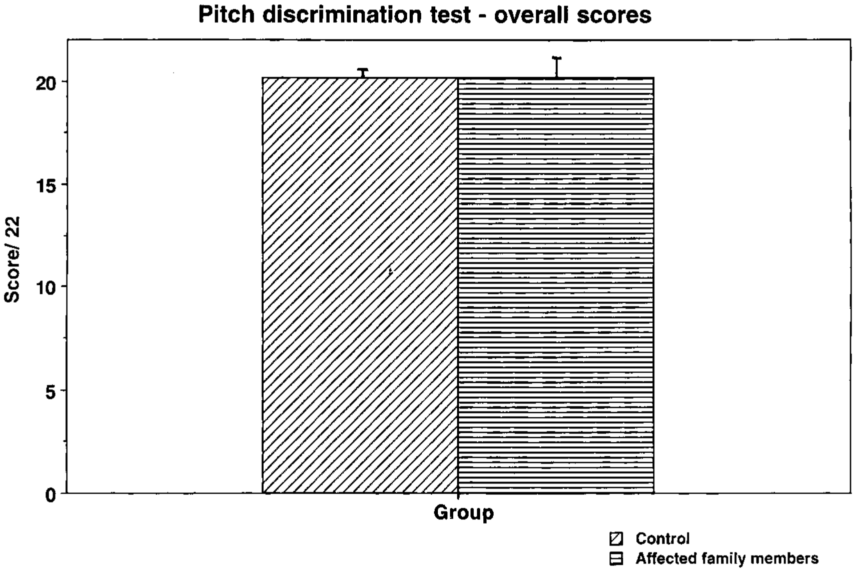

FIG. 1. Performance on pitch-discrimination task.

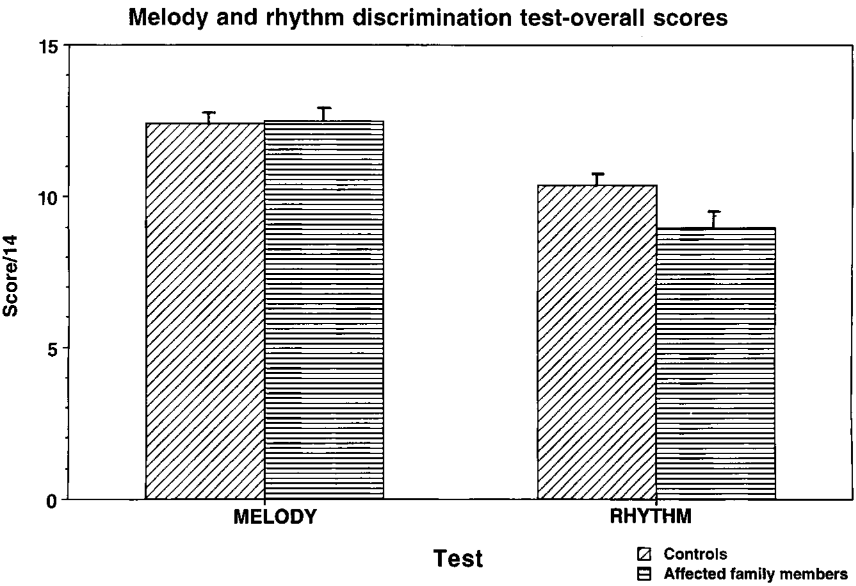

FIG. 2. Performance on melody- and rhythm-discrimination tasks.

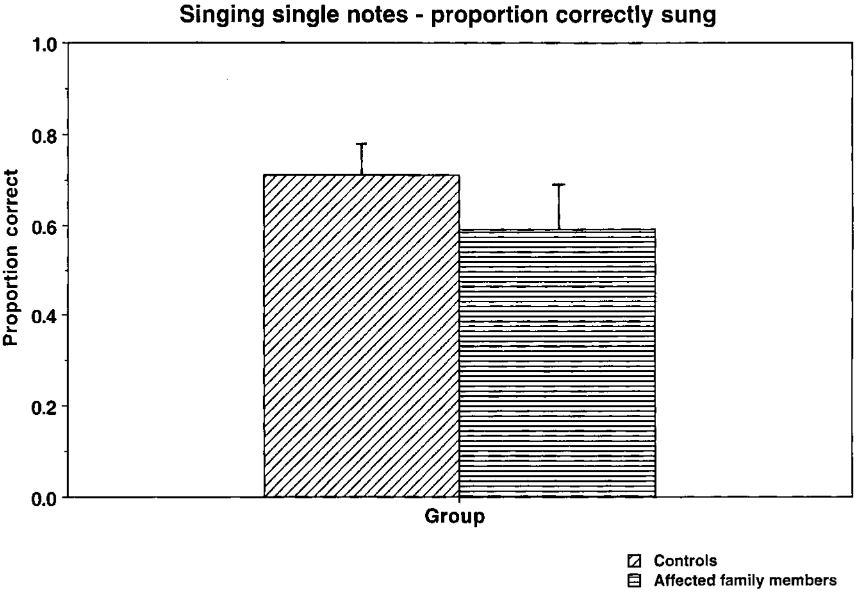

FIG. 3. Proportion of single notes sung correctly.

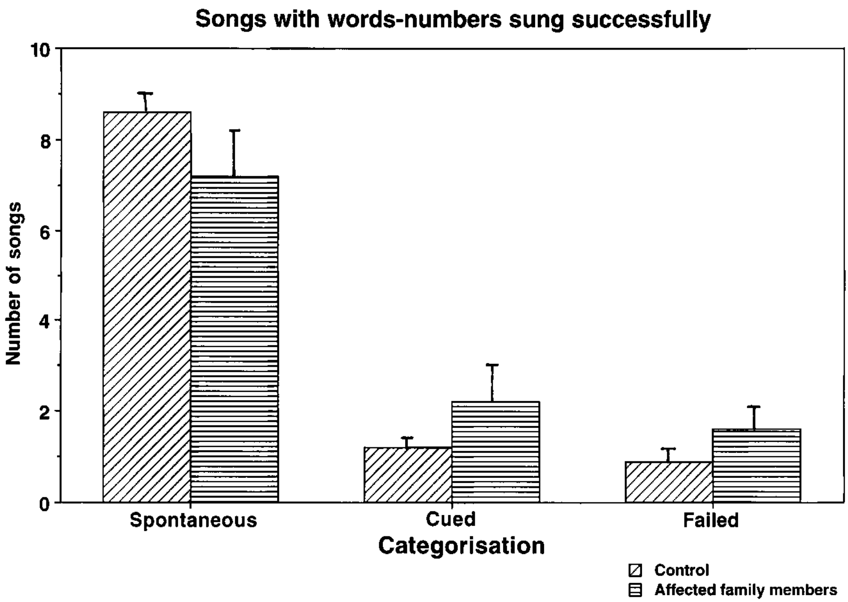

FIG. 4. Number of songs with words sung spontaneously or with a cue.

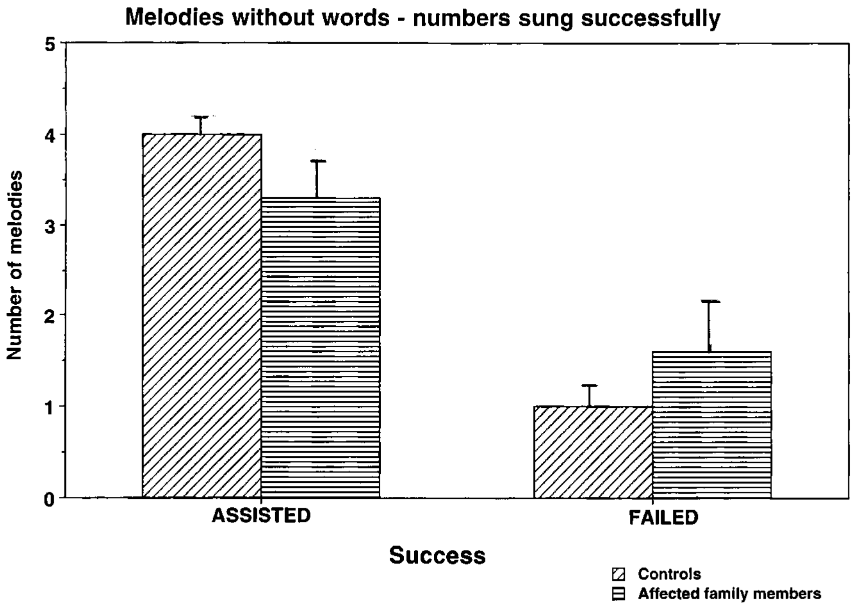

FIG. 5. Number of melodies without words sung successfully with a cue or not sung.

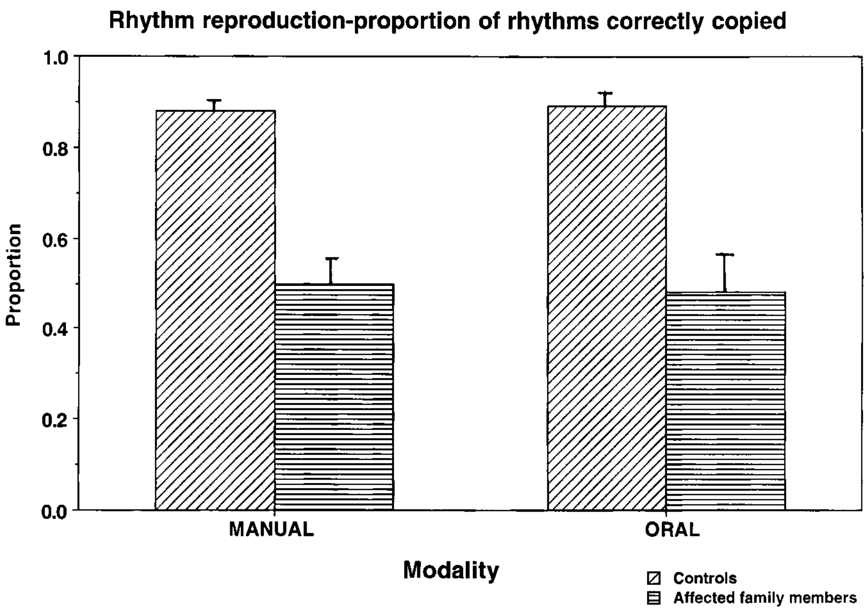

FIG. 6. Number of rhythms produced correctly, manual and oral conditions.

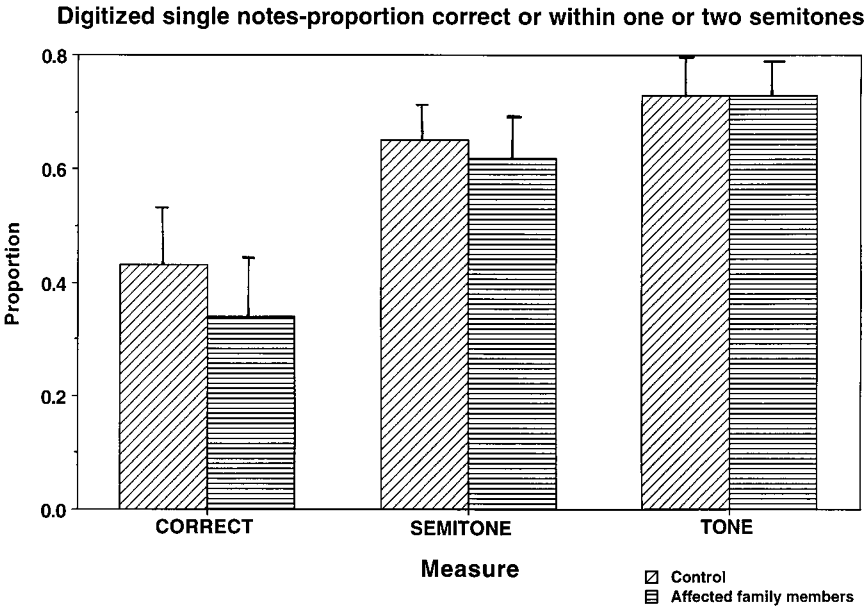

FIG. 7. Digitized single note singing: number of notes sung correctly or to within one or two semitones.

Similar results were obtained for proportion of notes correctly reproduced per rhythm (Modality: $F = 10.18$, $df = 1$, $p = .006$). One-way ANOVAs revealed that the affected family members performed significantly more poorly than the controls on the manual rhythms ($F = 4.55$, $df = 16$, $p < .05$) but not on the oral rhythms.

DISCUSSION

The affected family members were not impaired on any tasks involving musical intonation, but they were impaired on tasks involving the perception and production of rhythm.

On the music-perception tasks, affected family members were not impaired in pitch or melody discrimination, but they discriminated rhythms significantly less well than the controls. This is in agreement with the findings of Tallal et al. (1991), who observed that children with SLI performed more poorly on tasks involving perception of rapid timing.

On the music-production tasks, again, affected family members had no problems with any of the intonation tasks. The singing of single notes, and the singing of notes in melodies with or without words, whether scored by raters or digitized, were all performed as well by affected family members as by controls. Affected family members were no different on measures of number of songs sung, number of notes sung per song, and accuracy of notes

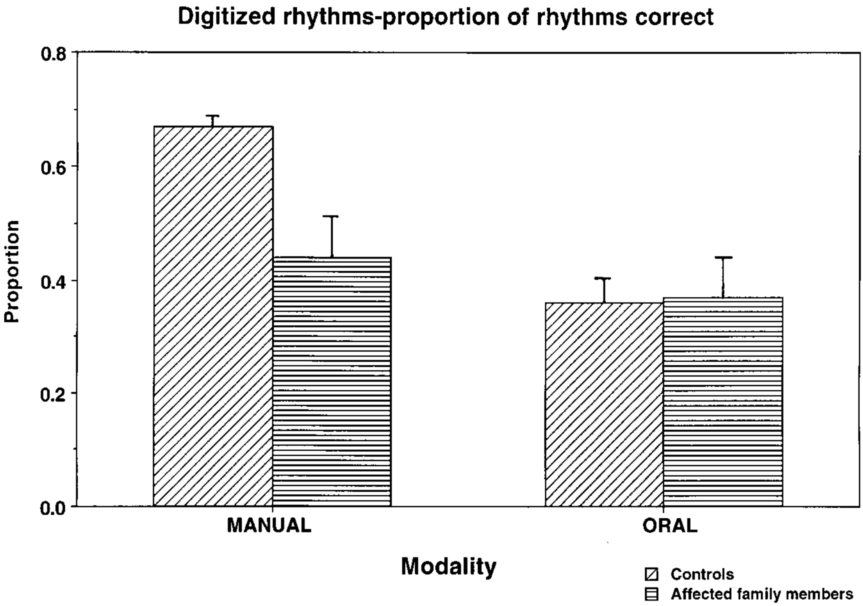

FIG. 8. Digitized rhythm production: number of rhythms correct, manual and oral conditions.

sung. Hence, affected family members have no problems with control of pitch in a musical context.

However, the affected family members were impaired on production of rhythms. When the productions were rated by the experimenter, an effect was found for both tapped (manual) and spoken (vocal) rhythms. However, when the productions were digitized, an effect was found for manual rhythms only (see Fig. 8).

Although the agreement between the two methods of scoring was extremely high (90%), it will be seen from a comparison of Figs. 6 and 7 that all subjects performed less well according to the acoustic analysis than according to the rating data, and this difference was particularly marked for the vocal rhythms. In the acoustic analysis, the placement of each spoken note can vary more than placement of the manual taps because the start of a note can be taken to be either the onset of the stop consonant or the onset of voicing. A rater will be consistent in this determination, whereas acoustic analysis may not. In addition, some positions in a rhythmic sequence will be shortened and others lengthened (Repp, 1992). These will be perceived by a musically trained observer as exactly corresponding to the presented rhythm, but will not be digitized as such. The above considerations suggest that scoring based on the digitized rhythms was not necessarily more accurate than scoring based on the ratings.

The rhythm-production tests are not direct language measures and, in the case of the manual rhythms, require no oromotor coordination. Therefore, the impairment on these tests cannot be explained by either a language deficit or an oral praxis deficit. When taken together with the impaired discrimination of rhythms, this difference is best explained by a central deficit in the processing of timing. Various authors, including Hammond (1982) and Tallal et al. (1991), have suggested that the underlying deficit which leads to speech and language disturbance is one of control and processing of timing. These data support that conclusion in the domains of both production and perception. Tallal et al.'s (1991) suggestion that a pervasive difficulty with temporal processing underlies specific language impairment seems to be upheld.

However, the stimuli used here—rhythms with fairly long intervals between taps or notes—may not correspond closely to the stimuli tested by Tallal and her colleagues (1991). These investigators used very short interstimulus intervals (in the 100-ms range), whereas in the present study the rates were slower, and in the production task it was adjusted for each subject as appropriate, even to a rate of 1 beat/s or slower if the subject required. On the other hand, some of the shortest intervals presented in the rhythm tasks were of the same order of magnitude as that used in the Tallal studies. Errors were not classified according to whether they were on short or long taps because erroneous reproduction of a rhythm tended to be in completely disordered, with virtually no taps of the right length or in the right place. This, together with the flexibly slow presentation, suggests that the affected family members did not have problems only with the shortest taps or notes. The impairment they displayed on the rhythm tasks may thus be one of patterns of timing rather than of processing very fast stimuli, possibly reflecting the relative timing deficit in dyspraxia proposed by Miller (1989).

Recent data on the neural basis of the verbal and oral dyspraxia in the affected family members (Vargha-Khadem et al., 1998) provide evidence as well regarding the neural basis of their timing difficulties. Abnormalities were found bilaterally in the head of the caudate nucleus as well as in many motor-related areas of the left hemisphere; these included an area of functional underactivity in the supplementary motor area (SMA), the same area in which Halsband et al. (1993) found that lesions disrupted the production of rhythms. The abnormalities in the caudate nucleus are also of interest in relation to the finding by O'Boyle et al. (1996) that patients with Parkinson's disease are impaired in the accurate timing of manual tapping. The combined evidence suggests that many of the same abnormalities that provide a possible explanation of the affected family members' dyspraxia symptoms could also help explain their impairment in temporal processing.

The current findings and those presented elsewhere (Alcock, 1995; Alcock et al., 1994, submitted) indicate that affected family members are deficient in a variety of verbal and nonverbal functions, including rhythm perception and production, production of complex single and combined oral movements, speech articulation, and a wide variety of linguistic processes. What might these deficits have in common? One possibility is a sequencing deficit combined with a fine-grained timing deficit. The former alone is unlikely to account for the difficulty they have in producing simultaneous oral movements, while the latter alone is a poor explanation for both the difficulty they have in producing sequences of oral movements and the deficits found here in fairly long, slow rhythms. Together, a timing and a sequencing could account for the impairment on combined movements of both types as well their rhythm deficits and their speech articulation problems. Articulation difficulty may reflect an impairment not only in oromotor control but also in phonological representation. Finally, subjects who have difficulty in producing phonologically accurate speech, and potentially in perceiving it, may then rationalize their grammatical production to maximize comprehension while compensating for the difficulty they have in producing and perceiving morphemes of low phonetic substance (Leonard, 1989).

Although an explanation in terms of a common underlying deficit is appealing, caution must always be exercised in attempting to account for one deficit in terms of another, and this is particularly the case when the deficits cover such a wide range. It is thus possible, instead, that in the affected members of the KE family several primary deficits coexist, each related to a different structural or functional abnormality among the several that have now been identified in these individuals.

REFERENCES

Alcock, K. J. (1995). *Motor dysphasia—A comparative study*. Ph.D. thesis, Oxford University.

Alcock, K. J., Passingham, R. E., Watkins, K., & Vargha-Khadem, F. (2000a). Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. *Brain and Language,* **75,** 17–33, doi:10.1006/brln.2000.2322.

Alcock, K. J., Wade, D., Anslow, P., & Passingham, R. E. (2000b). Pitch and timing abilities in adult left hemisphere dysphasic and right hemisphere damaged subjects. *Brain and Language*, **75,** 47–65, doi:10.1006/brln.2000.2324.

Botez, M. I., & Wertheim, N. (1959). Expressive aphasia and amusia following a right frontal lesion in a right-handed man. *Brain,* **82,** 186–202.

Damasio, A., Lima, A., & Damasio, M. (1975). Nervous function after right hemispherectomy. *Neurology,* **25,** 89–93.

Fisher, S. E., Vargha-Khadem, F., Watkins, K. E., Monaco, A. P., & Pembrey, M. E. (1998). Localisation of a gene implicated in a severe speech and language disorder. *Nature Genetics,* **18,** 168–170.

Halsband, U., Ito, N., Tanji, J., & Freund, H. J. (1993). The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain,* **116,** 243–266.

Hammond, G. R. (1982). Hemispheric differences in temporal resolution. *Brain and Cognition,* **1,** 95–118.

Leonard, L. B. (1989). Language learnability and specific language impairment in children. *Applied Psycholinguistics,* **10,** 179–202.

Miller, N. (1989). Apraxia of speech. In C. Code (Ed.), *The characteristics of aphasia*. London: Taylor & Francis. Pp. 130–154.

O'Boyle, D. J., Freeman, J. S., & Cody, F. W. (1996). The accuracy and precision of timing of self-paced, repetitive movements in subjects with Parkinson's disease. *Brain,* **119,** 51–70.

Penhune, V. B., Zatorre, R. J., & Evans, A. C. (1998). Cerebellar contributions to motor timing: A PET study of auditory and visual rhythm reproduction. *Journal of Cognitive Neuroscience,* **10,** 752–765.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain,* **113,** 1185–1205.

Phillips, D. P., & Farmer, M. E. (1990). Acquired word deafness, and the temporal grain of sound representation in the primary auditory-cortex. *Behavioural Brain Research,* **40,** 85–94.

Repp, B. H. (1992). Probing the cognitive representation of musical time. *Cognition,* **44,** 241–281.

Robin, D. A., Tranel, D., & Damasio, H. (1990). Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain and Language,* **39,** 539–555.

Samson, S., & Zatorre, R. J. (1991). Recognition memory for text and melody of songs after unilateral temporal lobe excision: Evidence for dual encoding. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **17,** 793–804.

Shankweiler, D. P. (1966). Effects of temporal lobe damage on the perception of dichotically presented melodies. *Journal of Comparative and Physiological Psychology,* **62,** 115.

Speedie, L. J., Wertman, E., Ta'ir, J., & Heilman, K. M. (1993). Disruption of automatic speech following a right basal ganglia lesion. *Neurology,* **43,** 1768–1774.

Stackhouse, J., & Snowling, M. (1992). Developmental verbal dyspraxia. II. A developmental perspective on two case studies. *European Journal of Disorders of Communication,* **27,** 35–54.

Tallal, P., Ross, R., & Curtiss, S. (1989). Familial aggregation in specific language impairment. *Journal of Speech and Hearing Disorders,* **54,** 167–173.

Tallal, P., Sainburg, R. L., & Jernigan, T. (1991). The neuropathology of developmental dysphasia: Behavioural, morphological and physiological evidence for a pervasive temporal processing disorder. *Reading and Writing,* **3,** 363–377.

Tzeng O. J. L., & Wang, W. S. Y. (1984). Search for a common neurocognitive mechanism for language and movements. *American Journal of Physiology,* **246,** R904–R911.

Vargha-Khadem, F., Watkins, K., Alcock, K., Fletcher, P., & Passingham, R. (1995). Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America,* **92,** 930–933.

Zatorre, R. J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia,* **23,** 31–41.