AN INVESTIGATION OF THE EFFECTIVENESS OF TWO METHODS OF STUDENT RESPONSE USING A TAPED PROGRAM OF PRACTICE MATERIALS FOR IMPROVING AURAL DISCRIMINATION

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

Rex J. Hewlett

1966

This is to certify that the

thesis entitled

AN INVESTIGATION OF THE EFFECTIVENESS OF TWO METHODS OF STUDENT RESPONSE USING A TAPED PROGRAM OF PRACTICE MATERIALS FOR INPROVING AURAL DISCRIMINATION

presented by

Rex J Hewlett

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Music

Merrell L. Harbaugh
Major professor

Date July 26, 1966

ABSTRACT

AN INVESTIGATION OF THE EFFECTIVENESS OF TWO METHODS OF STUDENT RESPONSE USING A TAPED PROGRAM OF PRACTICE MATERIALS FOR IMPROVING AURAL DISCRIMINATION

by Rex J. Hewlett

In the fall of 1965, an experimental study was conducted with part of the freshman theory classes at Central Michigan University. Purposes of the study were to:

1. Determine if improvement in aural discrimination would be greater if the practice response to an aural stimulus was played rather than written.
2. Learn if sight-singing progress would be greater if the dictation practice response mode were played rather than written.
3. Study the relationship of these two types of practice on improvement in error detection type of dictation.
4. Determine the effect of previous keyboard training on the two practice methods.
5. Examine the relationship of the response mode to high and low beginning levels of aural perception.
6. Observe any interaction between scholastic aptitude and (a) achievement in aural perception, and (b) the practice response mode.

Procedures

1. A set of dictation exercises was developed, printed, and recorded on magnetic tape.
2. Dictation and sight-singing tests were devised to serve as pre- and post-tests.
3. The study population was tested and divided into two equated groups which were then randomly assigned to the two practice methods. The experimental group played all responses at the keyboard; the control group used the traditional written response.
4. The resulting statistics were analyzed by rank order correlations and Mann-Whitney U tests.
5. The 5 per cent level of confidence was established as the significance level for the study.

Results

1. No significant difference was found between the two practice methods for overall dictation improvement.
2. No significant difference in improvement in sight-singing resulted from the two methods of practice.
3. Significant difference in favor of the experimental group occurred in error detection dictation.
4. There was no significant relationship between the two methods of practice and previous keyboard experience in overall dictation improvement.

5. In error detection significantly greater improvement was shown by experimental group students with previous keyboard experience. No significant difference existed between practice methods for the non-pianists.
6. Pre-test dictation score relationships with the improvement resulting from the two practice methods were not significant. A tendency was noted for control group improvement to be greater for those with low pre-test scores whereas experimental group improvement was greater for those with high pre-test scores.
7. Scholastic aptitude was a significant factor in dictation improvement for the control group but not for the experimental group.

Conclusions

1. The played response is a useful ear-training practice technique and should be added to the several existing practice methods.
2. Students who have piano background may find the played response very helpful, especially in the area of error detection.
3. Further study of the relationships between extreme initial levels of perception and the method of response was recommended.

4. The relationship between scholastic aptitude and the two methods of practice seemed worthy of further study.
5. The need for more sophisticated equipment for individual and laboratory classroom use was recognized.

AN INVESTIGATION OF THE EFFECTIVENESS OF TWO METHODS OF STUDENT RESPONSE USING A TAPED PROGRAM OF PRACTICE MATERIALS FOR IMPROVING AURAL DISCRIMINATION

By

Rex J. Hewlett

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Music

1966

ACKNOWLEDGMENTS

I am indebted to the members of the guidance committee: Dr. Richard Klausli, Dr. Merrell Sherburn, Dr. Robert Sidnell, and Dr. William Sur, for valuable assistance in this study.

Special acknowledgment is due to Dr. Sherburn, who gave invaluable assistance in the development of the experimental ideas and the programed materials, and to Dr. Sidnell for his help in the statistical design and preparation of the final report.

I am grateful to Dr. Sur, chairman of the committee, for continued encouragement and helpfulness throughout the work before and during the research period.

TABLE OF CONTENTS

		Page
ACKNOWLEDGMENTS		ii
LIST OF TABLES		iv
LIST OF FIGURES		v
Chapter		
I.	INTRODUCTION	1
	Experimental Hypotheses The Experiment Importance of the Study	
II.	ORGANIZATION OF THE STUDY	14
	Development of the Program Materials Development of the Practice Materials Sampling Procedure Experimental Controls Other Controls Evaluation Procedure Student Procedure Problems of the Study	
III.	RESULTS AND ANALYSIS OF EXPERIMENTAL DATA.	36
	Hypotheses Summary	
IV.	CONCLUSIONS, RECOMMENDATIONS AND SUMMARY	55
	Implications for Teaching Recommendations for Further Study Summary	
BIBLIOGRAPHY		68
APPENDICES		
A.	PRACTICE MATERIAL OF THE STUDY	71
B.	DICTATION AND SIGHT-SINGING TESTS	117

LIST OF TABLES

Table		Page
1.	Series A-Order of Interval Presentation. . .	19
2.	Series B-Order of Presentation of Time Patterns.	20
3.	Series C Program	21
4.	Control Factor Scores of Study Population . .	24
5.	Mean Scores of the Total Population	25
6.	Mean Scores of the Experimental and Control Groups	25
7.	Student Practice Procedure	31
8.	Pre-test and Post-test Overall Dictation Scores	38
9.	Pre-test and Post-test Error Detection Scores.	39
10.	Pre-test and Post-test Written Dictation Scores	41
11.	Dictation Pre-test and Post-test Scores According to Keyboard Training Groups . .	43
12.	Error Detection Gain Scores for Pianists and Non-pianists	45
13.	Dictation Gain Scores for Students with High and Low Dictation Pre-test Scores . . .	47
14.	Pre-test and Post-test Sight-singing Scores .	49
15.	Comparison of Dictation Gain and A.C.T. Scores	51
16.	Dictation Gain Scores for Students with High and Low A.C.T. Scores	52
17.	Pre-test, Post-test Rhythm Scores.	53

LIST OF FIGURES

Figure		Page
1.	Relationship of per cent gain scores in dictation to high and low dictation pre-test scores	62

CHAPTER I

INTRODUCTION

The development of latent sensitivity to musical sound has long been accepted as fundamental in the training of a musician. Schools of music have provided for this development through the music theory curriculum, which has traditionally included classes in written theory, sight-singing and ear-training. Contemporary pedagogical practice often combines these areas into one integrated course in musicianship.

In such integrated theory courses, Murphy recognizes six basic areas which should be considered: (1) writing, (2) playing, (3) reading, (4) listening, (5) analyzing, and (6) creating. He stresses that "listening is undeniably the sole basis of musical activity."[1]

The aspect of musical listening which involves discrimination between pitch and rhythm--i.e., ear-training, has been the subject of considerable attention which in recent years has resulted in a number of new training approaches. Such attention to this aspect of musical growth is understandable if one considers that, in addition to its necessity for adequate musical performance, listening

---

[1]Howard A. Murphy, Teaching Musicianship (New York: Coleman-Ross Company, Inc., 1962).

is of prime importance in all aspects of music teaching. The private teacher as well as the conductor of musical organizations must have an acutely developed sense of aural discrimination.

The traditional classroom ear-training process in which the teacher plays exercises in melodic, rhythmic, or harmonic dictation at the keyboard while the student listens and then attempts to notate what he has heard, is a slow and laborious procedure. The results are often questionable, particularly for the upper and lower extremes of the class. The gifted students are forced to drill in areas in which they may already be capable of satisfactory work; the students with little facility in dictation may receive insufficient practice for achievement of satisfactory results.

An essential largely missing from this traditional dictation procedure is the element of immediate reinforcement or feedback. The importance of such reinforcement in the learning process has long been recognized and it is the basis for contemporary programed instruction techniques.

In discussing response and reinforcement, Lumsdaine writes:

> The function of feedback or correction/confirmation following the student's response is closely related to the role that the response itself is conceived to have, and also to the degree of prompting prior to responding. . . . The first function is that of reinforcement in the sense of reward, presumed to strengthen the benefits of immediately preceding responding. The second function is that of providing prompts or cues for an additional <u>implicit response</u>, which may supercede, repeat, or fill in, at a covert

> level, the immediately preceding response depending on whether that response was, respectively, in error, correct, or lacking.[1]

This second function of reinforcement is that which, in addition to rewarding the student with the knowledge of the response correctness, actually operates as a supplemental informational source. It might be considered "directional" reinforcement which in the case of musical stimulus-response procedure would involve (a) information about an incorrect response, whether it was too high or too low, too long or too short, and (b) a "negative reward" for incorrect responses so that practice of errors would not be continued. In their generalizations of the reinforcement theory, Lysaught and Williams call this kind of reinforcement "differential reinforcement." Some of their conclusions are:

> 1. An individual learns or changes the way he acts by observing the consequences of his actions.
>
> 2. Consequences that strengthen the likelihood of repetition of an act are called reinforcements.
>
> 3. The more quickly reinforcement follows the desired performance, the more likely the behavior will be repeated.
>
> . . . . . .
>
> 7. The learning behavior of a student can be developed or shaped gradually by differential reinforcement--that is, by reinforcing those behaviors which should be repeated and by withholding reinforcement following undesired acts.[2]

---

[1] A. A. Lumsdaine (ed.), Student Response in Programmed Instruction (Washington, D. C.: National Academy of Sciences, National Research Council, Publication 943), Chapter 31.

[2] Jerome P. Lysaught and Clarence M. Williams, A Guide to Programmed Instruction (New York: John Wiley and Sons, Inc., 1963), Chapter I.

A possible implication from the above ideas is that, in the development of aural perception, the reinforcement might be stronger if the response was one which could be heard and compared aurally with the stimulus sounds. In this process the student would be bypassing the notation in favor of the actual sounds. This could conceivably result in an extinguishing process if the student were motivated, through hearing a distasteful sound, to correct his erroneous response. Similarly this kind of overt, sounded response could have a built-in prompting value through the recognition of a too high or too low response.

While the written response can also indicate, through comparison with the correct notation, whether the response is too high or too low, the inherent learning which could prevent a similar error on subsequent occasions, may not be as strongly present as it might be in the sounded response. This possibility would seem to be worthy of further consideration.

Considerable study has been directed toward learning the importance of the time lapse between the response and the knowledge of the correctness of that response. Ihrke comments on this aspect of reinforcement as it relates to the response mode as follows:

> It is important to point out that the write-in response does not provide immediate or even accurate feedback, since it requires at least ten seconds to write and compare with the correct model. If this visual comparison indicates that an error has been made, it is still not known whether the student heard correctly and wrote an error, or whether he

> heard incorrectly and wrote the misinformation correctly. Another problem . . . concerns the student's capability in reading the correctness or incorrectness of the response . . . If the response consists of writing a check mark in a box, the comparison with the correct answer is simple. But if he is asked to HEAR the difference between his erroneous response and the correct one, this may be considerably more complicated, and require of him a proficiency and ability which he at that moment does not possess.[1]

The search for improved ear-training techniques which could provide needed reinforcement has resulted in a variety of self-help systems. One of these is the use of the tape recording and disc recording for the presentation of stimuli for individual practice in dictation. This technique has been accepted as a useful tool in many schools of music.

As early as 1949 studies using tape recorded materials for ear-training practice were conducted by Cookson at Northwestern University. His comparisons of students practicing with conventional classroom ear-training procedures and those using tape recorded drill materials indicated the potential value of the latter technique.[2]

Sherburn prepared an extensive set of taped materials and developed effective procedural techniques which, over

---

[1]Walter R. Ihrke, "Trends in Music Education Research," Council for Research in Music Education, Bulletin No. 2 (Urbana, Illinois: University of Illinois, 1964), p. 30.

[2]Frank B. Cookson, Recordings and Self-tutoring (Cleveland, Ohio: The Brush Development Company, 1949).

a period of years, have been proven to be valuable in the improvement of aural discrimination. The practical value of these teaching devices is clearly demonstrated by his use of language laboratory facilities to provide practice opportunities for large numbers of music students.[1]

Spohn discovered that melodic dictation students who used tape recorded teaching materials for individual practice showed a 22.65 per cent greater decrease in errors in dictation than those in the typical classroom situation.[2]

In a similar study Carlsen found significant differences in favor of the group using the taped, programed materials and in a "side-effect" study he observed the progress of the two groups for improvement in sight-singing.[3] While the statistics indicated that the use of his materials did not significantly affect the improvement of sight-singing, subsequent evaluations by the theory teachers did show a favorable relationship for the experimental group. Further study of this problem was recommended.

Such studies were instrumental in establishing the programed self-help materials as valid instruments of

---

[1]Merrell L. Sherburn, "Music in the Language Laboratory," Music Educators Journal, January, 1964, 109-110.

[2]Charles L. Spohn, "An Exploration in the Use of Recorded Teaching Material to Develop Aural Comprehension in College Music Classes" (unpublished Ph.D. dissertation, Ohio State University, 1959).

[3]James C. Carlsen, "An Investigation of Programmed Learning in Melodic Dictation by Means of a Teaching Machine" (unpublished Ph.D. dissertation, Northwestern University, 1963).

learning. Other research has been directed toward discovering effective methods of presenting stimuli to the student and the effectiveness of various modes of student response. To this end a number of electronic devices have been developed which may prove useful in presenting aural practice stimuli and in providing immediate knowledge of response correctness. Some of these are adaptations of the tape recorder; others combine the tape recorder with other devices in more intricate, though not always practical, arrangements.

An intonation training device was used by McQuerrey in an attempt to improve sensitivity to interval intonation.[1]

Ihrke describes experimental studies in which student responses to notated rhythms were matched with an inaudible signal on a two-track tape. Immediate reinforcement was provided by flashing lights when errors occurred.[2]

Spohn carried out a four-pronged study of stimulus and response mode combinations. Combinations used were: (1) visual stimuli-written response, (2) visual stimuli-oral response, (3) aural stimuli-written response, (4) aural stimuli-oral response. The results of this study indicate

---

[1]Lawrence H. McQuerrey, "The Improvement of Sensitivity to Interval Intonation Through Training With A Mechanical Apparatus" (unpublished Ph.D. dissertation, Indiana University, 1957).

[2]Walter R. Ihrke, "Automated Music Training," Council of Research in Music Education Bulletin No. 2 (Urbana, Illinois: University of Illinois, 1964), p. 6.

that the student response mode is sometimes more important than the method of presenting the stimulus. While all four of the combinations were determined to be effective methods of practice, Spohn concluded that different methods may be helpful to different people. Similarly it appeared that one method was best for a certain task while a different one would be more effective for a second area of learning.[1]

If this diversity of usable practice techniques is desirable, it would seem logical that research studies be directed toward finding: (1) as many practical techniques as possible which could be helpful to students in their attempts to improve their aural perception, and (2) the techniques which are most helpful for particular students working with a specific musical element.

One technique which has been proposed as a useful method for individual practice is a played response to an aural stimulus. In considering the played response as a subject for possible research, several questions arise.

1. Might there be strong reinforcement of learning from hearing the response as it is played, and from the immediacy of the response?

Spohn concluded that a sung response is an effective method of improving aural comprehension; perhaps the

---

[1]Charles L. Spohn, "An Evaluation of Two Methods Using Magnetic Tape Recordings for Programmed Instruction in the Elemental Materials of Music" (final project report, Project 1407, Research Foundation, Ohio State University, 1964).

effectiveness stems from the actual hearing of the response rather than simply seeing it; perhaps temporal proximity is a factor since knowledge for the correctness of the response is possible within the response action itself.

2. Will the definite physical activity involved in the played response provide a keyboard image which could assist in the learning process?

Langsford, investigating the effects of varying amounts of practice on improvement of aural comprehension, observed the superior results for pianists and theorized that

> Knowledge of the piano keyboard and its associative value with certain pitches--i.e., mental imagery of the keyboard while taking melodic dictation may be of considerable value to some students of melodic dictation.[1]

If this theory is valid it leads logically to the conclusion that a played keyboard response would provide this helpful keyboard image.

3. Will such an image prove helpful in sight-singing progress?

4. Will such practice be practical for students without previous experience at the keyboard?

5. Will the student hear his incorrect response and will hearing it enable him to be more accurate in subsequent dictation situations?

6. Will written dictation suffer from a program consisting of keyboard practice exclusively?

---

[1]Harry Langsford, "An Experimental Study of Effect of Practice Upon Improvement in Melodic Dictation" (unpublished Ph.D. dissertation, Michigan State University, 1959).

7. Is the keyboard response technique a practical one?

While the electronic devices mentioned previously as being capable of providing immediacy and accuracy of reinforcement are still in the developmental stage and not readily available, the piano is already in general use in the keyboard phase of theory classes. Its practical value in terms of general use and availability is obvious; the technique of procedure would have to be worked out.

In consideration of the importance of the above questions it seemed advisable to devise an experimental study to: (a) develop a technique of keyboard response practice and (b) investigate the usefulness of such a technique.

Statement of the Problem

Can a programed course in dictation, using tape recorded stimuli and played responses at the keyboard, along with only a minimal basic practice in written responses, be more effective in teaching aural discrimination than the same programed materials using only written response techniques?

Specific Questions for Investigation

1. Will a played response develop aural discrimination as effectively as a written response?
2. Will sight-singing be favorably affected if the student practices a played response instead of a written response?

Experimental Hypotheses

In the form of hypotheses the questions are:

1. Improvement in aural discrimination will be significantly greater when the practice response to an aural stimulus is played than when it is written.

2. Sight-singing progress will be significantly greater when the dictation practice response mode is played than when it is written.

Other Questions

Additional questions which will be considered in this study are:

1. Will the played response be a more effective practice method for error detection or corrective dictation than the written response method?

2. What effect will previous keyboard training have on the use of the keyboard response method? Can a non-pianist make effective use of the keyboard response technique?

3. Does the response mode exert any significant influence on the improvement of aural discrimination for students whose pre-test dictation scores are at the extreme high and low ends of the scale?

4. What interaction exists between scholastic aptitude and (a) achievement in aural perception, (b) response method in training aural perception?

The Experiment

An experimental design was developed in which two groups of students were established and pre-tested. The sampling process is described in Chapter II. The experimental group practiced with a prepared set of tape recorded exercises. Their responses to the recorded stimuli were played back at the piano keyboard. Feedback was possible

in two ways: (a) an aural comparison of the played response with a second presentation of the stimuli, and (b) comparison of the response with a printed sheet of correct responses. The control group used the same set of materials but constructed written responses in the traditional manner of dictation practice. Reinforcement for this group consisted of a comparison of the response with a printed sheet of correct answers. At the conclusion of the experiment the two groups were tested again and compared statistically for improvement in aural comprehension.

A complete description of experimental procedures and the resulting data appears in Chapters II and III.

Importance of the Study

Presently accepted techniques of practice in ear-training have not completely satisfied the needs of all students who wish to study music. It has been recognized that a generalizing of results of a particular practice technique, while perhaps desirable, is not completely possible and that research is needed to discover the particular methods of practice which will be valid and practical for students with particular needs. Lumsdaine emphasizes the importance of studying the response factor in learning when he states that

> Student responses and the techniques by which they may be suitably controlled for effective learning are of both practical and theoretical interest.

> Their role in any systematic account of learning is of central importance.[1]

This study might answer the question of whether the played response would satisfy this kind of need in a practical manner.

---

[1]A. A. Lumsdaine, "The Analysis of Student Response as a Factor in Instruction," Student Response in Programmed Instruction (Washington, D. C.: National Academy of Sciences-National Research Council, 1961), Chapter I.

CHAPTER II

ORGANIZATION OF THE STUDY

Development of the Program Materials

The basic concepts which guided the development of the materials for this experiment were:

1. Aural perception can be effectively strengthened through the use of tape programed practice materials.
2. Practice exercises for improving aural pitch perception can be presented effectively through iso-rhythmic tone patterns.
3. Rhythm perception can be as effectively improved through exercises presented in a melodic context as in exercises presented on a single repeated tone.
4. Error detection is a valid technique for evaluating the development of aural discrimination.

That taped programed materials can be effective practice tools has been demonstrated by several observers as noted in Chapter I.

Presenting exercises in pitch perception through iso-rhythmic tone patterns is a well established practice. Various sight-singing and dictation manuals using this type of exercise have proven their usefulness through actual practice. At least one experimenter, Reeves, in 1954 discovered that there is no significant difference in progress

in development of pitch perception if training material is of the drill type instead of actual melodies.[1]

Admittedly the placing of time patterns in a melodic framework would not present a purely rhythmic exercise and might tend to distort statistics in this area; however, there were other considerations for this study. Since the ultimate goal of this type of research is the development and testing of materials and techniques which might be useful in actual teaching situations, the recognition of the strong practical interrelationship between rhythm and melody justified the use of rhythm within some melodic pattern. This interrelationship seems more evident when listening to isolated rhythmic structures than when hearing isolated tone patterns.

Other considerations were the difficulty of playing repeated tones at the keyboard, particularly for the non-pianist, and the tendency of repeated single tones to create unfavorable attitudes in listeners. In an informal study made earlier by the writer, students were found to react more favorably to each of several other techniques of presenting rhythm than to repeated single tone presentations.

---

[1]William Reeves, "An Exploratory Study of Two Sets of Theories of Guthrie and Wheeler as they Relate to the Development of Instrumental Musicianship" (unpublished Ph. D. dissertation, University of California, 1954).

Murphy notes that error detection or "corrective listening" dictation is a useful type of practice.

> It is organized on the principle that one of the most practical uses of aural sensitivity is the detection of errors in music heard. This is a perfectly reasonable assumption which has many implications for teaching. Obviously conductors constantly are required to do this very thing in rehearsals, but it also applies with equal validity to any musician who is called upon to make decisions involving the eye versus the ear, as for example, the teacher of applied music . . .[1]

Development of the Practice Materials

It was decided that the study should be limited to three areas: pitch, rhythm, and complete melodic dictation. Three series of exercises, one for each of the areas to be studied, were written and recorded on magnetic tape. Each series consisted of fifteen tapes approximately twenty minutes in length.

Series A exercises were iso-rhythmic tone patterns. The exercises progressed from the intervals believed to be least difficult to hear, within a tonal context, to the more difficult ones and from short three-note patterns to a maximum of eleven notes.

Research studies regarding the difficulty of hearing intervals are contradictory. Spohn observed that a certain order of difficulty prevails as follows: P8, M2, m2,

---

[1]Howard A. Murphy, Teaching Musicianship (New York: Coleman-Ross Company, Inc., 1962), p. 74.

M3, P4, P5, M6, M7, m3, A4, m7, m6, progressing from least difficult to most difficult.[1]

In an earlier study, Ortmann concluded that training in interval hearing should proceed from small to large intervals.[2] More recently, Jeffries discovered that drilling on intervals of increasing difficulty did not produce better results than from a randomized order.[3] In the absence of established guidelines it was arbitrarily decided that the order of presentation of intervals would be: M2, m2, M3, m3, P5, P8, P4, M6, M7, m6, m7, A4.

Consideration for the non-pianists and the need for basic simplicity prompted the decision to use the major keys of C, F, G, and D for the majority of exercises. It was thus possible to avoid difficult and awkward fingerings and at the same time guard against the possible limiting effects of an exclusive C major orientation. It will be noted in Tables 1, 2, and 3 that the keys of Eb and A major and a, c, d, and e minor were used in the higher

---

[1]Charles L. Spohn, "An Evaluation of Two Methods Using Magnetic Tape Recordings for Programmed Instruction in the Elemental Materials of Music," Project 1047 (Research Foundation, Ohio State University, 1964), p. 13.

[2]Otto Ortmann, "Some Tonal Determinants of Melodic Memory," Journal of Educational Psychology, XXIV (1933), 454-467.

[3]Thomas Broad Jeffries, "The Effects of Order of Presentation and Knowledge of Results on the Aural Recognition of Melodic Intervals" (unpublished Ph.D. dissertation, University of California at Los Angeles, 1965).

tape numbers of the series when students had gained greater facility at the keyboard. Table 1 shows the progressive nature of interval use, clefs and keys in Series A exercises.

Series B exercises were limited to practice in rhythm perception. As noted earlier, all exercises were presented within the same melodic framework; the first five tones of the major scale were selected as a suitable grouping.

Table 2 shows the rhythmic progress schedule for Series B exercises as well as the clef, key and meter involved in each exercise. It will be noted that only the last three tapes of the series made use of a specific meter; the others were simply rhythmic patterns without concern for pulse accents.

Practice exercises in Series C combined pitch and rhythm in complete melodic dictation. All materials used in Series A and B were included in Series C exercises. The program for Series C is shown in Table 3. The entire set of exercises which made up the three taped series is shown in Appendix A.

Sampling Procedure

The population of this study consisted of students in first year music theory classes at Central Michigan University. Three laboratory sections were involved, two of which were taught by the experimenter; the other by a

TABLE 1

SERIES A-ORDER OF INTERVAL PRESENTATION

TAPE	CLEF	KEY	M2	m2	M3	m3	P5	P8	P4	M6	M7	m6	m7	A4
1	G*	C	X	X										
2	G	G,D	X	X										
3	G	C,F	X	X	X	X								
4	G/F	G,C	X	X	X	X								
5	G/F	F,D	X	X	X	X	X	X						
6	G/F	F,G	X	X	X	X	X	X						
7	G	C,F	X	X	X	X	X	X	X					
8	F	G,D	X	X	X	X	X	X	X					
9	G	C	X	X	X	X	X	X	X	X				
10	G	F	X	X	X	X	X	X	X	X				
11	G	G	X	X	X	X	X	X	X	X	X			
12	G	a	X	X	X	X	X	X	X	X	X	X	X	X
13	G/F	d	X	X	X	X	X	X	X	X	X	X	X	X
14	G/F	F,e	X	X	X	X	X	X	X	X	X	X	X	X
15	G	c,Eb	X	X	X	X	X	X	X	X	X	X	X	X

*The G or treble clef and F or Bass clef are the only clefs used throughout the study.

TABLE 2

SERIES B-ORDER OF PRESENTATION OF TIME PATTERNS

Time Patterns	Tape Number														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Whole units	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Simple division			x	x	x	x			x	x	x	x	x		
Dotted quarters					x	x					x	x	x		
Syncopation					x	x		x						x	x
Compound division							x	x						x	x
Unit subdivision									x	x	x	x	x		
Dotted sixteenths											x	x	x		
3-4 meter													x		
6-8 meter														x	
9-8 meter															x
Key	G	C	D	A	G	C	F	G	F	G	D	F	G	A	F
Clef	g	f	g	f	g	f	g	g	g	g	g	g	f	g	f

TABLE 3

SERIES C PROGRAM

Intervals & Rhythm	Tape Number														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
M2	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
m2	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
M3			x	x	x	x	x	x	x	x	x	x	x	x	x
m3			x	x	x	x	x	x	x	x	x	x	x	x	x
P5					x	x	x	x	x	x	x	x	x	x	x
P8						x	x	x	x	x	x	x	x	x	x
P4			x	x	x	x	x	x	x	x	x	x	x	x	x
M6			x				x	x	x	x	x		x	x	x
M7														x	x
m6			x				x	x	x	x	x		x	x	x
m7									x	x		x		x	x
A4						x	x			x	x	x	x		x
Whole units	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Simple division				x	x	x	x			x	x		x	x	
Dotted quarter notes						x	x			x	x		x	x	
Syncopation						x	x			x		x		x	x
Compound division								x	x			x			x
Sub-division														x	
2-4 meter														x	
3-4 meter		x			x		x				x				
4-4 meter	x		x	x		x				x			x		
6-8 meter								x	x			x			
Key	C	F	C	G	D	F	C	G	C	a	d	g	C	F	d

member of the theory staff of the University. A fourth section, designated as an honors section because of high scores received on a theory and ear-training placement examination given at Central Michigan University, was not involved in this study.

Experimental Controls

Experimental controls for the study included:

1. Scores attained on a pre-test in aural perception designed especially for this study.
2. Scores of a pre-test in sight-singing ability designed for this study.
3. Percentile scores of the American College Test given to all entering students at Central Michigan University.
4. Results of a piano placement audition given by the keyboard staff of the Department of Music including information concerning previous keyboard study.

The piano rating was an ordinal rating based on actual performance and information concerning previous study and the solos and studies involved in this experience. The information was used in this study primarily to define whether a student had previously had any study in keyboard or not.

The original group balance included nearly an equal number of students from each piano rating group in each of the study groups.

From these scores and ratings, two essentially equivalent groups were formed and assigned randomly to the two practice response modes. However, during the course of

the experiment the number of subjects was reduced to twenty-nine due to normal attrition and certain other factors which are discussed more completely later in this chapter. Table 4 shows the tabulation of each of the scored factors for all students who completed the study. Table 5 shows the mean scores for each of the scored factors and Table 6 gives the comparison of mean scores for the two groups.

Other Controls

The teaching of the three laboratory sections was planned so that learning experiences would be similar in each section. Basically these experiences include rhythmic, melodic and harmonic dictation, sight-singing, keyboard practice and acquaintance with some of the standard music literature. During the course of this experiment the practice in rhythmic and melodic dictation was eliminated from the laboratory class periods. Since students of both the experimental and control groups were scattered through all three sections, it was felt that any effects of teaching differences would be minimized for the final comparisons.

The possibility of Hawthorne effect influencing the final comparisons of the two groups in this experiment was greatly reduced due to the fact that students at Central Michigan University had not previously had an organized tape practice program in ear-training; thus one group was not considered more experimental than the other. Further

TABLE 4

CONTROL FACTOR SCORES OF STUDY POPULATION

Dictation Errors	Sight-singing Errors	A.C.T. %-iles	Piano Ratings*
93	80	07	2
114	60	81	3
104	73	16	4
117	74	16	3
109	79	21	4
111	59	10	3
118	48	33	3
104	49	10	1
99	49	16	4
109	36	77	2
96	80	69	2
102	66	65	3
106	66	95	1
83	99	16	2
82	54	33	2
80	28	44	4
98	59	65	1
111	44	16	3
97	89	96	1
99	76	74	1
88	51	39	1
76	33	65	2
108	32	33	1
108	83	44	1
114	66	83	4
117	63	29	2
119	69	88	3
118	76	33	4
114	88	55	1

*Piano rating: 4 indicates highest rating, 1 the lowest.

TABLE 5

MEAN SCORES OF THE TOTAL POPULATION

Scored Factors	Mean	SD
Dictation pre-test (errors)	103.24	12.12
Sight-singing pre-test (errors)	63.08	18.62
A.C.T. percentile scores	45.83	28.42
Piano placement ratings	2.34	---

TABLE 6

MEAN SCORES OF THE EXPERIMENTAL AND CONTROL GROUPS

Scored Factors	Control		Experimental	
	Mean	SD	Mean	SD
Dictation pre-test (errors)	104.50	9.40	101.93	13.50
Sight-singing pre-test (errors)	65.57	17.01	60.73	19.32
A.C.T. percentiles	38.00	31.15	53.13	23.59
Piano placement	2.64	---	2.06	---

any suggestion that one group might be expected to improve more than the other was carefully avoided.

Students were informed that they would not be graded on the pre-test or post-test results but that the tapes covered a part of the course work over which they would be tested, apart from the experiment itself.

Evaluation Procedure

It was necessary to determine the student levels of aural perception at the beginning and the completion of the study. Tests in sight-singing and dictation which could serve as both pre-test and post-test were essential. The features desired in the tests included:

1. A dictation test composed of two sections
   a. written dictation
   b. error detection (including not only the detection of errors but the correction of these errors)
2. A sight-singing test of sufficient length to provide a fair test of sight-singing ability
3. Items of sufficient difficulty in each test to provide a challenge to all students in post-testing
4. Enough material of an elementary nature in each test to make possible an adequate pre-test judgment

None of the available standardized tests seemed to meet all of these qualifications and a special test was devised in each of the areas which would satisfy the testing needs of the study.

Dictation Pre- and Post-Test

Items 1, 2, and 3 involved iso-rhythmic error detection. The test sheet contained examples which the students were asked to compare with a played version in which certain errors were inserted. Incorrect notes were to be circled by the student and the correct note or accidental inserted. If the student failed to circle an incorrect note, or if he circled a correct note, a two-point error was recorded. A one-point error occurred if the incorrect note was circled but not corrected.

Item 4 was an exercise in rhythmic error detection. In this exercise the beat unit was considered an entity, and students simply placed a check mark below any beat which was heard to be different from the printed version. Failure to detect the differences correctly resulted in a one-point error.

Items 5 and 6 were combined pitch and rhythm examples in which a printed exercise differed from a played version. Students were asked to circle each incorrect note or beat unit and indicate with a letter P or R whether the difference was one of pitch or rhythm. A one-point error was awarded for each failure to detect either a pitch or rhythm difference.

The remainder of the test consisted of eight exercises, four in iso-rhythmic dictation and four of complete melodic dictation. These eight exercises required a written

response. Each pitch was valued at one point, and on the last four each beat unit was valued at one point.

The entire dictation test is shown in Appendix B.

Sight-Singing Pre- and Post-Test

Items 1 and 2 were short scale number patterns to be sung at any comfortable pitch level selected by the student. Items 3, 4, and 5 were eight measure melodies selected because they contained problems covered in the programed materials of the study, particularly those in the area of pitch.

Students were allowed approximately fifteen seconds to look over each test item and then were asked to sing the exercise through without stops if possible. Each student's effort was recorded on tape for greater accuracy in scoring. A pitch or beat unit which was sung incorrectly was recorded as a one-point error. If both pitch and rhythm were wrong a two-point error was scored. A stop or hesitation sufficient to seriously break the rhythm was recorded as a rhythm error. At certain pre-selected points a student who had lost the tonality was stopped and given the pertinent pitch, then asked to continue. A one-point error was recorded for each stop of this nature.

The complete sight-singing test is shown in Appendix B.

Validity and Reliability of Tests

The assertion of validity for each of the tests is based on the fact that the materials included in each test were not different from those covered by the programed materials of the study and are traditional musical considerations in the first year theory course in ear-training and sight-singing at Central Michigan University and other institutions. In addition the responses required in the tests are not different from those outlined in the objectives of this study.

Both of the pre-tests were subjected to statistical measures to check their reliability. Using the split-half method, the reliability coefficients for the two tests were found to be .75 for the dictation test and .74 for the sight-singing test.

Student Procedure

All students were instructed to listen to each of the forty-five tapes twice. Each exercise was played twice on the tapes and an answer sheet was provided with each tape, providing opportunity for reinforcement in the learning process.

The control group listened and constructed a written response for each exercise. An immediate check with the answer sheet was recommended. If the student wished, he could play the exercise a third time while watching the answer sheet for combined aural and visual reinforcement.

Those in the experimental group listened, then played their response at the piano keyboard. A check with the answer sheet after the second hearing and response, or repeating the response along with the taped repeat of the exercise provided reinforcement possibilities for this group.

It was recommended that these students determine the key of the exercises on a particular tape and play the scale and tonic arpeggio in that key before starting their practice; this would tend to lessen technical difficulties for the non-pianists. In Series A tone pattern exercises, students were asked to speak note names or scale numbers while playing their responses. For the rhythm exercises of Series B, the keyboard group was advised to use some system of counting note values to help develop the image of the note played. Table 7 shows the complete practice procedure for both groups.

The pre-tests were given during the fourth week of the semester and the students began use of the tapes in the fifth week. The delayed beginning of the study was planned to allow time for the students to become oriented to theory classes, procedure of dictation, sight-singing, the keyboard, and notation of music.

Ideally, the dictation post-test would have been given immediately after each student completed the tapes, but it was not possible to schedule them in this way, and

TABLE 7

STUDENT PRACTICE PROCEDURE

Taped Stimuli	Control Group	Experimental Group
Exercise played	Listen	Listen
Silence	Write exercise	Play exercise
Exercise repeated	Listen	Listen OR play
Silence	Finish writing	Play
	Check response	Check response
Stop the tape for further checking on response OR proceed to the next exercise.		

the test was given after all twenty-nine students had completed their practice. It was felt that no great contamination resulted since all but one student finished less than one week before the post-test date. This one person completed practice elven days before the test was given. The sight-singing post-test was given as soon as each person could be scheduled after he had completed the tapes.

Printed instructions concerning the practice procedure and use of the equipment supplemented an oral presentation of this material in each laboratory section by the experimenter. A suggested schedule for tape use was provided for each student. This was intended to help maintain an orderly pace which would be more meaningful than a less regular pace; however, this did not restrict individuals from progressing at a faster pace if desired.

Equipment and Facilities

Three complete sets of tapes were prepared for the study. All were recorded at seven and one-half inches per second, using an Ampex tape recorder, Model 601 and an Altec microphone. These tapes were placed in the Music Department record library; librarians checked out tapes and listening equipment to the students.

Listening equipment available for this study included:

4	Wollensak tape recorders	Model T-1515
1	Revere tape recorder	Model T-1100
1	Revere tape recorder	Model T-202
1	Robyn tape recorder	Model 200
2	Lafayette tape decks	
12	Koss headphone sets	Model SP-3

Practice facilities available for the study included forty-five practice rooms in the Music building, each of which is equipped with a piano and electric outlets. Additional facilities which could be used by the control group were the tape decks in the listening area of the Music building.

Problems of the Study

Certain difficulties encountered during the course of the experiment should be noted. The first of these involved a problem of motivation in one of the laboratory sections which resulted in a number of students not completing the experiment. A second group of students was lost between semesters as individuals dropped out or failed the course. One student was hospitalized and unable to

complete the semester. These factors combined to reduce the study population to twenty-nine students, fifteen in the experimental group and fourteen in the control group.

These changes in the structure of the two groups, though not substantially altering the dictation and sight-singing mean scores, caused a greater variance in balance in keyboard ratings and American College Test percentile score means.

The scores of the two groups as finally constituted were subjected to Mann-Whitney U tests with the following results:[1]

1. No significant difference existed for the dictation pre-test. U = 100.
2. No significant difference existed for the sight-singing pre-test. U = 93.50.
3. The difference for A.C.T. scores was significant in favor of the experimental group. U = 65.50.

As a result of the structural changes of the population, the decision was made to use non-parametric statistical measures to test the experimental data obtained from the study. These tests are useful when certain assumptions about the study population are not possible. Siegel observes that

---

[1]The Mann-Whitney U test is one of the most powerful of the non-parametric tests. The statistic U has a known sampling distribution from which it is possible to determine probability of occurrence.

> These newer 'distribution free' or non-parametric techniques result in conclusions which require fewer qualifications. Having used one of them, we may say that 'regardless of the shape of the population(s), we may conclude that . . .'.[1]

and further:

> . . . because behavioral scientists rarely achieve the sort of measurement which permits the meaningful use of parametric tests, non-parametric statistical tests deserve an increasingly prominent role in research in the behavioral sciences.[2]

A second problem which should be noted was the impracticability of effective monitoring of the individual practice sessions. Laboratory facilities featuring electronic keyboards and monitoring devices, had they been available, could have reduced any contamination of data which may have resulted from this lack of monitoring. Lacking this equipment, an effort was made to control the practice sessions by making regular spot checks during the times when the practice tapes were available to the students, to determine if they were carrying out the proper practice procedures. It was possible to do this without interrupting the students in their practice sessions.

It was not possible to control certain variables which could have had some effect on the results of the study.

---

[1]Sidney Siegel, Nonparametric Statistics for the Behavioral Sciences (New York: McGraw-Hill Book Co., Inc., 1956), p. 3.

[2]Ibid., p. 31.

Such factors as musical maturation which involved student participation in other classes, private lessons, performing organizations and recital attendance along with their general physical and mental growth, are possible contaminants.

All of the above mentioned problems are recognized as impracticable to control and for purposes of replication and generalization their control may not be essential. In connection with this matter, Scott and Wertheimer make this observation:

> In laboratory studies, the particular functional relationship observed may be in large part a product of the special, artificial conditions of the experiment. Such a restriction in generality is likely to be particularly serious in those 'hold constant' experiments in which control of extraneous variables artifically places them at levels that would hardly be encountered in the subject's normal activity. A given experiment may show conclusively a relationship between the independent and dependent variables, which, however, is demonstrated only under such peculiar circumstances that the result is uninteresting, except as a curiosity.[1]

---

[1]William A. Scott and Michael Wertheimer, Introduction to Psychological Research (New York: John Wiley and Sons, Inc., 1962).

CHAPTER III

RESULTS AND ANALYSIS OF EXPERIMENTAL DATA

At the conclusion of the practice period the information obtained from the experimental procedures was compiled and subjected to statistical analysis. The content of this chapter deals with the results of these analyses. The conclusions and recommendations are a part of Chapter IV.

Hypotheses

Hypothesis #1

> Improvement in aural discrimination will be significantly greater when the practice response to an aural stimulus is played than when it is written.

The first hypothesis was tested for significance of the per cent decrease in error of the 15 students in the experimental group and the 14 students in the control group. The most suitable non-parametric test for this is the Mann-Whitney U test. The value of U was found to equal 81.5. For groups of this size a U of 66 or less is necessary for significance at the 5 per cent level of confidence selected for use in this study.

Although there was a slightly greater mean improvement in the overall dictation scores by the experimental group

the difference was not significant and hypothesis #1 was rejected.

Table 8 shows the pre-test, post-test per cent gain scores and means for the two groups.

One of the additional questions raised in Chapter I in connection with dictation improvement was:

> Will the played response be a more effective practice method for error detection or corrective dictation than the written response method?

The error detection section of the dictation test was scored separately on both the pre-test and post-test. Computation of the U for corrective dictation gains of the two groups results in a U of 61. This difference is significant at the 5 per cent level of confidence. The conclusion is that the played response practice method is significantly more effective in reducing errors of corrective dictation than the written method. Question #1 would be answered in the affirmative.

The pre-test, post-test scores and gains of the two groups are shown in Table 9.

A related concern is a consideration of the effectiveness of the two methods of practice in improvement of written response. Will the control group produce significantly higher gains than the experimental group when tested only on written dictation?

The Mann-Whitney U for constructed dictation gain is U = 97. Since a U of 66 or smaller is required for

TABLE 8

PRE-TEST AND POST-TEST DICTATION SCORES

Experimental Group			Control Group		
Pre	Post	% Gain	Pre	Post	% Gain
76	41	46	83	64	23
80	29	60	93	63	32
82	72	12	96	80	17
88	53	40	99	79	20
97	77	21	102	77	25
98	100	-2	104	82	21
99	66	33	104	97	7
108	75	31	106	67	37
108	86	20	109	82	25
111	96	13	109	91	17
114	80	30	111	73	34
114	83	27	114	69	41
117	89	24	117	91	22
118	84	29	118	75	37
119	82	31			
Mean 101.93	Mean 74.20	Mean 27.67	Mean 104.50	Mean 77.86	Mean 25.57

TABLE 9

PRE-TEST AND POST-TEST ERROR DETECTION SCORES

Experimental Group			Control Group		
Pre	Post	% Gain	Pre	Post	% Gain
25	9	64	25	19	24
18	5	72	23	14	39
29	22	24	30	26	13
32	15	53	33	34	-3
28	14	50	30	26	13
33	35	-6	33	27	18
37	23	38	38	46	-21
33	29	12	32	21	34
40	32	20	35	27	23
30	29	3	33	29	12
39	22	44	32	33	-3
36	22	39	29	23	21
31	20	3	30	28	7
42	30	29	37	24	35
38	25	34			
Mean 32.73	Mean 22.80	Mean 31.93	Mean 31.43	Mean 26.93	Mean 15.14

significance at the 5 per cent level the conclusion is that there is no significant difference between the two groups in improvement in written dictation.

Table 10 shows the scores and mean gains of the two groups for the written dictation section of the pre- and post-tests.

Question #2 in Chapter I asked:

> What effect will previous keyboard training have on the use of a keyboard response practice method? Will persons who have had no previous keyboard experience be able to make effective use of the keyboard response technique?

In order to test this question the dictation gain scores were separated into two groups, those with no keyboard background (piano rating 1) and those with varying degrees of previous training (piano ratings 2, 3, and 4). No attempt was made to consider the amount or quality of training except that all those in the first group, hereafter called non-pianists, had had no previous keyboard training and those in the second group, designated as pianists, had previously had at least a semester of piano instruction of some sort.

The U test was applied to the scores of the nine non-pianists and twenty pianists. The result showed that U = 83. For groups of this size a U of 54 or less is required for significance at the 5 per cent level. The conclusion is that non-pianists do not differ significantly from pianists in overall dictation gain scores when considered without regard to practice method.

TABLE 10

PRE-TEST AND POST-TEST WRITTEN DICTATION SCORES

Experimental Group			Control Group		
Pre	Post	% Gain	Pre	Post	% Gain
51	32	37	58	45	23
62	24	61	79	49	30
53	50	6	66	54	18
56	38	32	66	45	32
69	63	9	72	51	30
65	65	0	71	55	23
62	43	30	66	51	23
75	46	39	74	46	38
68	54	21	74	55	25
81	67	17	76	62	18
75	58	23	79	40	50
78	61	22	85	46	46
86	59	32	88	63	30
76	54	30	81	51	37
81	57	30			
Mean 69.20	Mean 51.40	Mean 25.93	Mean 73.93	Mean 50.93	Mean 30.21

Although seven of the nine non-pianists of the study were students of the experimental group, a comparison was made of dictation gain scores of control group non-pianists and pianists. The U for this group was 10.5 which is not significant at the 5 per cent level and it appears that there is no significant difference between levels of piano experience for control group students in the matter of dictation gain.

A similar comparison was made between the two piano groups of the experimental group. The result was a U of 25. For this group this does not reach the 5 per cent level of confidence and there was no apparent difference in dictation gains as a result of variance in piano background in this group.

Table 11 shows the scores of per cent gain in dictation for the two degrees of piano background for the control, experimental and combined groups.

The significantly higher gains of the experimental group in error detection reported earlier in this chapter introduce questions concerning the relationship between the piano background and error detection gains for the two methods of practice.

For those with no piano background the Mann-Whitney U test of the gain difference between the experimental and control groups results in a U of 3 which has a probability of .167 and is not significant at the 5 per cent level

TABLE 11

DICTATION PRE-TEST AND POST-TEST SCORES ACCORDING TO KEYBOARD TRAINING GROUPS

Group	Experimental			Control		
	Pre	Post	% Gain	Pre	Post	% Gain
Non-Piano	88	53	40	104	97	6
	97	77	21	106	67	37
	98	100	-2			
	99	66	33			
	108	75	31			
	108	86	20			
	114	83	27			
	Mean	Mean	Mean	Mean	Mean	Mean
	101.7	77.14	19.86	105.0	82.0	21.50
Piano	76	41	46	83	64	23
	80	29	60	93	63	32
	82	72	12	96	80	17
	111	96	13	99	79	20
	114	80	30	102	77	25
	117	89	24	104	82	21
	118	84	29	109	82	25
	119	82	31	109	91	17
				111	73	34
				114	69	41
				117	91	22
				118	75	37
	Mean	Mean	Mean	Mean	Mean	Mean
	102.1	71.62	30.62	104.5	77.17	26.17

selected for this study. Error detection gains are not significantly different between the two practice methods for the students without previous keyboard training.

The comparison of the error detection gains of the experimental and control groups of students who had had previous keyboard training involved the scores of twenty students, eight in the experimental group and twelve in the control group. The mean per cent gain is 34.00 for the experimental group and 17.08 for the control group. The U test for these scores results in a U of 22 which is significant at the 5 per cent level of confidence. It appears that among students who have had some keyboard training, those who used the played practice response method made greater gains in error detection than those who used the written practice technique.

Table 12 shows the error detection gain comparisons for the two piano groups. Such a great difference in mean scores for non-pianists without attaining the 5 per cent level is possible because of the numbers involved; two were in the control group and seven in the experimental group.

Question #3 of Chapter I was:

> Does the response mode exert any significant influence on the improvement in aural discrimination for students whose pre-test dictation scores are at the extreme high and low ends of the scale?

TABLE 12

ERROR DETECTION GAIN SCORES FOR PIANISTS AND NON-PIANISTS

Group	Experimental			Control		
	Pre	Post	% Gain	Pre	Post	% Gain
Non-Piano	32	15	53	38	46	- 21
	28	14	50	32	21	34
	36	22	39			
	37	23	38			
	40	32	20			
	33	29	12			
	33	35	- 6			
	Mean	Mean	Mean	Mean	Mean	Mean
	34.14	24.28	29.43	35.50	38.50	6.50
Piano	25	9	64	25	19	24
	18	5	72	23	14	39
	29	22	24	30	26	13
	30	29	3	33	34	- 3
	39	22	44	30	26	13
	31	30	3	33	27	18
	42	30	28	35	27	23
	38	25	34	33	29	18
				32	33	- 3
				29	23	21
				30	28	7
				37	24	35
	Mean	Mean	Mean	Mean	Mean	Mean
	31.50	21.50	34.00	30.83	25.83	17.08

To test this question the four highest scores of the dictation pre-test were compared with the five lowest scores of the same test for the experimental group. The inclusion of a greater number of low scores in this comparison was necessary as the result of tied scores at the low end of the scale.

The U test was applied to the gain scores with a resulting U of 5 which has a probability of .143. This does not reach the level of significance decided as necessary for this study and the answer to question #3 is that there is no significant difference between gains of high and low pre-test dictation score students in the experimental group.

A similar application of the U test on the four highest and four lowest pre-test dictation scores in the control group shows a U of 2 which has a probability of .57, almost reaching the 5 per cent level of confidence. It should be noted that the per cent gain for the experimental group was greater for students who had high pre-test scores in dictation whereas in the control group those who had the lowest pre-test dictation scores made the greatest gains.

Table 13 shows the high and low dictation pre-test scores and per cent gain scores for the experimental and control group.

TABLE 13

DICTATION GAIN SCORES FOR STUDENTS WITH HIGH AND LOW DICTATION PRE-TEST SCORES

Group	Experimental			Control		
	Pre	Post	% Gain	Pre	Post	% Gain
High Pre-Test	76	41	46	83	64	23
	80	29	60	93	63	32
	82	72	12	96	80	17
	88	53	40	99	79	20
	Mean 81.50	Mean 48.75	Mean 39.50	Mean 92.75	Mean 71.50	Mean 23.00
Low Pre-Test	114	80	30	111	73	34
	114	83	28	114	69	41
	117	89	24	117	91	22
	118	84	29	118	75	37
	119	82	31			
	Mean 116.40	Mean 83.60	Mean 28.40	Mean 115.0	Mean 77.0	Mean 33.50
	U = 5 p = .143			U = 2 p = .057		

Hypothesis #2

> Sight-singing progress will be significantly greater when the dictation practice response mode is played than when it is written.

The second hypothesis was also tested by means of the Mann-Whitney U test. The mean per cent gain scores for the two study groups were 30.87 for the experimental group and 28.29 for the control group. The value of U = 104.5 for these scores. Significant difference at the 5 per cent level would require a U of 66 or less. It appears that the two methods of practice show no significant difference in relation to sight-singing improvement and hypothesis #2 is rejected.

Table 14 shows the pre-test and post-test scores and means for the two groups in sight-singing.

Question #4 in Chapter I concerned the existence of interaction between scholastic aptitude and (a) achievement in aural perception and (b) the response method for training aural perception. Consideration of this problem was given by applying the following statistical measures:

1. Rank order correlation of dictation per cent gain scores with American College Test percentiles.
2. Mann-Whitney U test of dictation gains between experimental group students and A.C.T. percentiles.
3. Mann-Whitney U test of dictation gains between control group students and A.C.T. percentiles.

The rank order correlation for dictation gain scores and A.C.T. percentile scores produced an r = .308 which is

TABLE 14

PRE-TEST AND POST-TEST SIGHT-SINGING SCORES

Experimental Group			Control Group		
Pre	Post	% Gain	Pre	Post	% Gain
33	14	58	99	42	57
28	14	50	80	49	39
54	39	28	80	50	37
51	36	30	49	29	41
89	73	18	66	45	32
59	67	-13	73	26	64
76	38	50	49	56	14
32	42	-31	66	47	29
83	40	52	36	44	-22
44	36	18	79	47	41
66	21	68	59	28	53
88	67	24	60	37	38
63	33	48	74	69	7
76	50	34	48	51	-6
69	49	29			
Mean 60.73	Mean 41.26	Mean 30.87	Mean 65.57	Mean 44.28	Mean 28.29

not significant at the 5 per cent level of confidence. It appears that the dictation gain is not related to the student A.C.T. score in any significant manner.

Table 15 shows the computation of the rank order correlation of A.C.T. percentile scores and dictation gain scores for the entire study population.

To check the relationship between A.C.T. scores and the experimental group dictation gains, the four highest and five lowest A.C.T. scores of this group were selected. The extra score in the low group was included as the result of tied scores in that group. The U test resulted in a U of 3.5 which has a probability of .075. The 5 per cent level of significance was not reached and the difference in gains between students of high and low scholastic aptitude was not considered significant.

Students of the control group who had the three highest and nine lowest A.C.T. percentile scores were compared. The larger number of low scores was used to maintain the same score range as used in the experimental group above. The U of 3 for these scores was significant at the 5 per cent level and it appears that a significant difference exists between dictation gains of high and low scholastic aptitude students who used the written dictation practice techniques. The students with high A.C.T. scores made greater gains than those with low A.C.T. scores.

TABLE 15

COMPARISON OF DICTATION GAIN AND A.C.T. SCORES

Per Cent Gain	A.C.T. Scores
46	65
60	44
12	33
40	39
21	96
- 2	65
33	74
31	33
20	44
13	16
30	83
27	55
24	29
29	33
31	88
23	16
32	7
17	69
20	16
25	65
21	16
7	10
37	95
25	77
17	21
34	10
41	81
22	16
37	33

r = .308

Table 16 shows the highest and lowest A.C.T. scores and dictation per cent gains of the experimental and control groups.

TABLE 16

DICTATION GAIN SCORES FOR STUDENTS WITH HIGH AND LOW A.C.T. SCORES

Experimental				Control			
A.C.T.	Pre	Post	% Gain	A.C.T.	Pre	Post	% Gain
96	97	77	21	95	106	67	37
88	119	82	31	81	114	69	39
83	114	80	30	77	109	82	25
74	99	66	33				
			Mean 28.75				Mean 33.70
33	82	72	12	33	118	75	36
33	108	75	31	21	109	91	17
33	118	84	29	16	83	64	23
29	117	89	24	16	99	79	20
16	111	96	13	16	104	82	21
				16	117	91	22
				10	104	97	7
				10	111	73	34
				7	93	63	32
			Mean 21.80				Mean 23.55
U = 3.5 p = .075				U = 3 p = .05			

The rhythmic factors of the dictation test were scored separately and the U test applied to the pre-test, post-test scores of the two practice method groups. The result was a U of 104. With U = 66 or less required for the 5 per cent level of confidence, the difference in scores of the two groups were not considered significant.

Table 17 shows the pre-test, post-test scores and per cent gains for the two main study groups.

TABLE 17

PRE-TEST POST-TEST RHYTHM SCORES

Experimental			Control		
Pre	Post	% Gain	Pre	Post	% Gain
21	6	15	15	10	33
20	6	70	19	12	37
14	12	14	14	10	28
10	4	60	20	17	15
8	8	0	20	13	35
14	7	50	19	17	11
15	7	53	19	19	0
25	15	40	21	2	90
12	18	-50	20	10	50
30	22	27	27	15	44
27	17	37	29	13	55
24	11	54	27	7	74
32	19	40	28	12	57
19	11	42	32	15	53
27	13	52			
Mean	Mean	Mean	Mean	Mean	Mean
19.20	11.73	37.33	22.14	12.28	41.57

Summary

The data of this chapter is summarized as follows:

1. Hypothesis #1 was rejected. No significant difference was noted between the two study groups in overall dictation improvement.

2. Hypothesis #2 was rejected. No significant difference was discovered between the two groups in sight-singing improvement.

3. Error detection improvement was significantly greater for the experimental group.

4. Written dictation improvement was not significantly different for the two practice methods.

5. No significant difference in dictation gain was noticed between groups as the result of piano background.

6. Non-pianists were not significantly different in error detection gains as a result of the practice technique used but pianists were different in favor of the experimental group.

7. Students who received high and low pre-test scores were not significantly different in dictation gains in the two groups although the control group showed a .057 probability of difference in favor of the low pre-test students.

8. Rank order correlation between the A.C.T. scores and dictation gain scores showed no significant relationship for the population as a whole.

9. The highest A.C.T. score students showed greater actual gains in dictation in both study groups. The control group difference was significant at the 5 per cent level and the experimental group at .075.

10. Improvement in rhythm dictation was not significantly different for the two practice method groups.

CHAPTER IV

CONCLUSIONS, RECOMMENDATIONS AND SUMMARY

The data which is presented in Chapter III provides considerable material for study. Although the two main hypotheses are rejected, a fact which at first might appear to minimize the value of the experiment, there are other observations which may have important implications for ear-training practice techniques. This chapter considers the implications and recommendations resulting from the data of Chapter III.

Experimental hypothesis #1 proposed that dictation improvement would be greater for the experimental group than for the control group. The data rejects this hypothesis at the 5 per cent level of confidence. The mean per cent gain for the control group was 25.57 and for the experimental group, 27.67. While this latter gain was not significantly higher than that of the control group, neither was it significantly lower. This leads to the conclusion that the played response practice method is at least as useful as the generally used written response method. Students who are having difficulty with written responses may find the played response a more suitable technique.

Before recommending the keyboard practice technique for inclusion in the ear-training program, it is essential to consider its effect on written dictation. Will written dictation be adversely affected by an absence of transfer of learning from the played response? A corollary to this is the question: Will error detection or "corrective" dictation show a greater improvement with the use of the played response method?

Approaching the first question it was found that the pre-test, post-test mean per cent gain of the control group was 30.21 compared with 25.93 for the experimental group. The statistical analysis again disclosed that the difference was not significant at the 5 per cent level. Stated differently, the statistics reject the idea that there was a significant difference in gain in written dictation as the result of the practice method. Here again the absence of difference has importance for the study, inasmuch as the students of the experimental group, although they did not practice with the written technique, were able to show almost the same amount of improvement in written dictation as those who practiced entirely with the written method.

A conclusion of this study might be that, assuming a basic amount of practice in the notation of music, further practice in written dictation is not essential to the development of a good aural-visual sensitivity.

Of greater importance to the pedagogy of ear-training, because of its obvious practical implications for the teaching and conducting of musical performers, is the discovery that in error detection dictation, the experimental group gain was significantly higher than the control group, when piano background was not a considered factor. The mean per cent gain for the control group was 15.14 and 31.93 for the experimental group. The Mann-Whitney U test showed the difference to be significantly higher for the played response method at the five per cent level of confidence.

These statistics indicate a definite superiority of the played response technique over the written response in the area of error detection. While the reasons for this difference are not clear, a possible explanation may be that the played response produces a greater continual involvement with sound itself, thus providing more actual practice in listening. The student who practices with the written response technique may be expending greater effort on notation than on actual listening.

Obviously there is need for further research in this matter. This study involves only melodic and rhythmic error detection; whether or not the superior results of this kind of practice will extend to harmonic dictation and error detection must still be determined.

Having established a potential usefulness for the keyboard practice method in general, the impact of

previous keyboard training on the effectiveness of such practice was considered. A study of the difference scores in dictation for students who had no keyboard background and for those with a semester or more of instruction revealed important results.

The data shows that the mean per cent gain scores of non-pianists in overall dictation were lower than those of the pianists when the experimental and control groups were considered separately. These differences, however, did not reach the required level of significance. Similarly, when the scores of the combined groups were considered, without regard for the practice method used, the difference resulting from previous training was not significant.

It appears that piano background is not a significant factor in dictation improvement when the dictation includes both written and corrective types. This is in seeming contradiction with the earlier conclusions of Langsford; however, the actual per cent gain scores do show the same kind of relationship as he found.[1] Perhaps a replication of this study extending the population size, would be in order.

A comparison of the effect of the two practice methods on the overall dictation scores of pianists and

---

[1]Harry Langsford, "An Experimental Study of Effect of Practice Upon Improvement in Melodic Dictation" (unpublished Ph.D. dissertation, Michigan State Unisvesity, 1959).

non-pianists showed the experimental group gains to be slightly higher than the control groups but these gains do not reach the 5 per cent level of significance. It can be assumed that the presence or absence of previous keyboard training does not greatly affect the usefulness of either of the two practice methods for overall dictation.

To further test the influence of keyboard training on the practice method results, its relationship with error detection improvement was investigated. The improvement shown by the students without previous keyboard experience was not significantly different between the control and experimental groups. These non-pianists seemed to perform equally well with either practice technique. Again the implication is that the played response method will be at least as effective as the usual written response practice even though the student has not had previous training at the keyboard.

One should be cognizant of the fact, however, that the mean scores in this comparison are greatly different; 29.43 per cent gain for the experimental group; 6.50 per cent for the control group. The rejection of significant difference at the 5 per cent level is influenced heavily by the number of scores involved. With seven scores in the experimental group and only two in the control group, and in view of the wide difference of scores, it seems

reasonable that further study involving larger populations should be carried out with non-pianists using the played practice response method.

Continuing the investigation of the effectiveness of the played response in relation to keyboard experience disclosed that among those with some previous keyboard training there was significant difference between the experimental and control groups in error detection. The experimental pianists' gains were significantly higher than those of the control group. This was a not unexpected result, since the experimental group as a whole achieved greater gains than the control group in error detection dictation.

Another question considered in relation to overall dictation gains was whether or not students who have extreme high or low level pre-test scores in dictation would benefit more from the keyboard practice method than from the written method.

A comparison of the control group per cent gain scores for the two levels indicated a difference which was significant at .057, almost reaching the level for this study. The mean difference was in favor of the students with the low pre-test scores. There appears to be an inverse relationship between amount of gain and the beginning level as indicated by the pre-test scores of this group.

For the experimental group a similar comparison indicated that students with high pre-test scores made greater mean gains than those with low pre-test scores, but the difference did not reach the 5 per cent level of confidence.

It is interesting to note that the actual score differences resulting from the two practice methods, though not considered significant for this study, favor opposite groups. The mean difference for the control group indicated that students with high pre-test scores who practice with written responses do not make as much progress as those with low pre-test scores. Conversely, the experimental group students who had the high pre-test scores made more progress than the low score group. Figure 1 shows these relationships.

As noted in the piano relationships earlier, and now in regard to beginning levels of aural discrimination, the experimental practice method seems to provide a technique for the student who is initially better equipped, which may enable him to achieve higher levels of perception than the written response practice technique. At the same time, it appears that the written practice method does not provide better help than the experimental method for the students who have low pre-test scores. Both methods appear to have definite areas of usefulness.

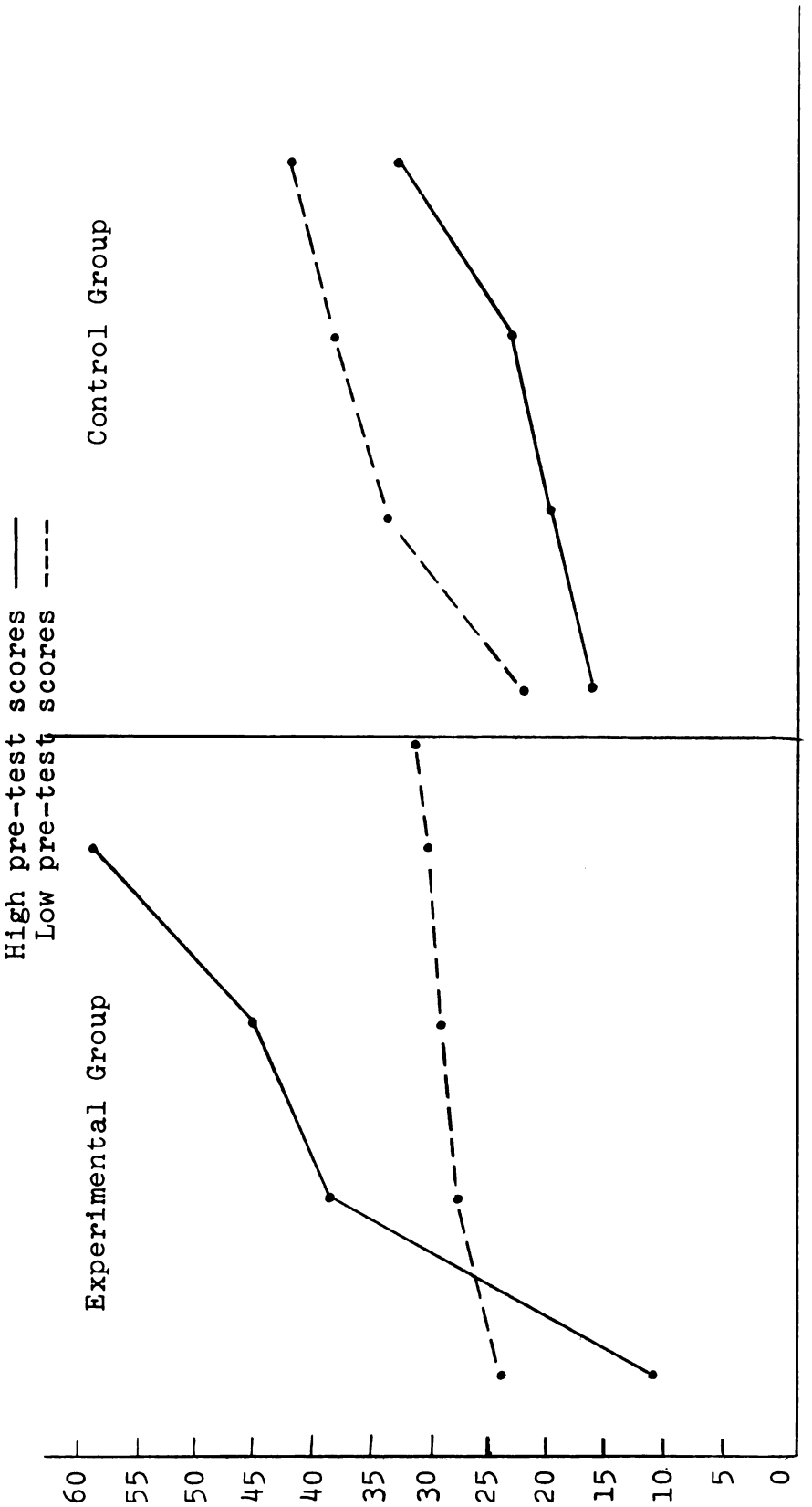


Figure 1.--Relationship of per cent gain scores in dictation to high and low dictation pre-test scores.

Experimental hypothesis #2 proposed that the experimental practice method would result in a superior improvement in sight-singing over the control method. The tests rejected this hypothesis at the 5 per cent level. Apparently there was little relationship between practice method and sight-singing improvement for this study.

The possibility is strong that other factors, impossible to control experimentally, were operative here. The amount of singing activity in which the student was involved--voice techniques classes, vocal ensembles, and any other such activity--could have a particular influence on the growth in sight-singing ability, especially for those who had not done any singing before their college enrollment. This factor varied not only between individuals but with individuals as they changed classes between semesters, making it impracticable to control. Any resulting contamination would be difficult to assess as to degree or direction.

Scholastic aptitude, as measured by the A.C.T. tests, was considered in relation to dictation gain for the two study groups. A rank difference correlation showed an $r = .308$ which is not significant for the study. Apparently there is little overall relation between dictation improvement and scholastic aptitude in this study.

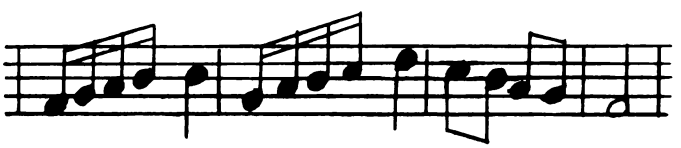
Previous studies have shown that the relationship between scholastic aptitude and musical ability is not

significant, but for the learning of musical skills such as notation, the relationship is thought to be positive. With this in mind it might seem logical to predict that the written method of practice might be more closely related to scholastic aptitude than the played response. This is exactly what is demonstrated by the data of this study. When comparing the highest and lowest A.C.T. scores with the two practice methods it was found that no significant difference existed within the experimental group. For the control group the difference was significant at the 5 per cent level in favor of the high A.C.T. scores. It should be noted, however, that the experimental level of significance was .075, indicating that certain other factors, such as learning the keyboard and counting, make the experimental method more than simply an expression of musical ability.

In the matter of rhythmic improvement no significant difference is demonstrated between the two practice methods. The mean per cent gain score of 37.33 for the experimental group is influenced by a negative gain of 50 per cent by one of the students. Both methods of practice appeared to be effective in improving the rhythmic element of aural perception.

A factor which has importance when rhythm is involved in the practice, whether rhythm practice alone or in combination with melody, is the time element. It was

discovered that when the rhythm gets as involved as simple division of the beat, it takes longer to notate an exercise than it does simply to play it back at the keyboard. This factor is multiplied as the rhythmic problems become more complicated. While a comprehensive time study was not included in the experiment, an example will illustrate the relative amounts of time involved in the two techniques. Exercise #14-6 of Series C requires approximately 15 seconds



to write in a simple copying situation; it can be played in 7 seconds at M.M. = 72. Other examples offer more striking evidence of the time saving element inherent in the played response. Within a specified time span, this could provide greatly increased opportunity for drill.

Implications for Teaching

The implications of this study for the teaching of aural comprehension are as follows:

1. Since the keyboard practice response has been shown to be a useful technique, it should be made available as an additional technique for practice in ear-training.

2. The played response appeared to be especially useful in improving error detection. It should be available to aid in this practical aspect of ear-training.

3. Pianists may find the played response a useful device for increasing their proficiency in ear-training. They should be encouraged to make use of the technique.

4. Additional techniques of keyboard practice may prove helpful and practical such as: (a) paired practice in which two persons send musical ideas back and forth at the keyboard, and (b) practice responses to taped harmonic dictation at the keyboard, cadences, chord progressions, choral phrases.

5. The use of electronic keyboard devices, some of which are already available, may prove practical for whole classrooms practicing responses to taped stimuli or to stimuli produced by the instructor at a master keyboard.

Recommendations for Further Study

As a result of the data obtained and the conclusions drawn from the data, the following recommendations for further study are suggested:

1. A study would be valuable to determine if the played response could be useful in harmonic dictation and the limits to which such dictation might be carried.

2. It would be useful to learn if the played response practice would be practical and effective on instruments other than keyboard instruments.

3. A study involving a larger population to check the influence of keyboard training on overall dictation improvement.

4. A three-pronged study using written, sung, and played responses to aural stimuli to compare the effectiveness of the three practice methods.

5. Further investigation of the apparent superiority of the keyboard response technique as a training device for students who achieve high levels of aural perception at the outset of their ear-training classes.

Summary

The findings of the study are summarized as follows:

1. The experimental method was found to be a useful device for dictation practice, though not significantly better than the written method for general dictation.

2. For error detection practice the keyboard response technique was more effective than the written method when considered without regard for piano background.

3. No significant difference between practice methods was observed for the non-pianist group although the actual mean gain score was higher for those in the experimental group.

4. Students who practiced exclusively with the written response did not show significantly greater improvement in written dictation than those who practiced only the played response.

5. Keyboard background is not a significant factor in overall dictation improvement. Non-pianists were able to achieve at about the same rate as the pianists.

6. In error detection, pianists had significantly higher gains when the response was played than when it was written. Non-pianists were not significantly different for either practice method.

7. High and low extremes of dictation pre-test scores were not significantly related to progress in dictation.

8. Sight-singing progress was neither favorably nor unfavorably affected by the experimental method of practice.

9. Scholastic aptitude is apparently more closely related to the written method than the played method, perhaps because of the skills required for notation.

10. Rhythm gains are not significantly different between practice method groups; both groups seemed to improve considerably using the taped practice materials.

11. Dictation drill is less time consuming when the played response is used than when a notated response is made.

BIBLIOGRAPHY

BIBLIOGRAPHY

Books

Cookson, Frank B. Recordings and Self-tutoring. Cleveland, Ohio: The Brush Development Company, 1949.

Fish, Arnold and Lloyd, Norman. Fundamentals of Sight-singing and Ear-training. New York: Dodd, Mead and Company, 1964.

Garrett, Henry E. Statistics in Psychology and Education. New York: Longmans, Green and Company, 1958.

Lindquist, Everett F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton-Mifflin Company, 1953.

Lumsdaine, Arthur A. (ed.). Student Response in Programmed Instruction. Washington, D. C.: National Academy of Sciences--National Research Council, 1961.

Lysaught, Jerome P. and Williams, Clarence M. A Guide to Programmed Instruction. New York: John Wiley and Sons, Inc., 1963.

Murphy, Howard A. Teaching Musicianship. New York: Coleman Ross Company, Inc., 1962.

Ottman, Robert W. Music for Sight-singing. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1956.

Scott, William A. and Wertheimer, Michael. Introduction to Psychological Research. New York: John Wiley and Sons, Inc., 1962.

Siegel, Sidney. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company, Inc., 1956.

Articles and Periodicals

Ihrke, Walter R. "Automated Music Training," Council of Research in Music Education, Bulletin No. 2 (Winter, 1964).

Ihrke, Walter R. "Automated Music Training," Journal of Research in Music Education, XI, No. 1 (Spring, 1963), 3-20.

Ortmann, Otto. "Some Tonal Determinants of Melodic Memory," Journal of Educational Psychology, XXIV (September, 1933).

Sherburn, Merrell L. "Music in the Language Laboratory," Music Educators Journal, January, 1964, 109-110.

Project Report

Spohn, Charles L. "An Evaluation of Two Methods Using Magnetic Tape Recordings for Programmed Instruction in the Elemental Materials of Music." Project Report, Project No. 1407, Research Foundation, Ohio State University, Columbus, Ohio, 1964.

Unpublished Material

Carlsen, James Caldwell. "An Investigation of Programmed Learning in Melodic Dictation by Means of a Teaching Machine." Unpublished Ph.D. dissertation, Northwestern University, Evanston, Illinois, 1963.

Jeffries, Thomas Broad. "The Effects of Order of Presentation and Knowledge of Results on the Aural Recognition of Melodic Intervals." Unpublished Ph.D. dissertation, University of California at Los Angeles, 1965.

Langsford, Harry. "An Experimental Study of Effect of Practice Upon Improvement in Melodic Dictation." Unpublished Ph.D. dissertation, Michigan State University, East Lansing, Michigan, 1959.

McQuerrey, Lawrence H. "The Improvement of Sensitivity to Interval Intonation Through Training with Mechanical Apparatus." Unpublished Ph.D. dissertation, Indiana University, Bloomington, Indiana, 1959.

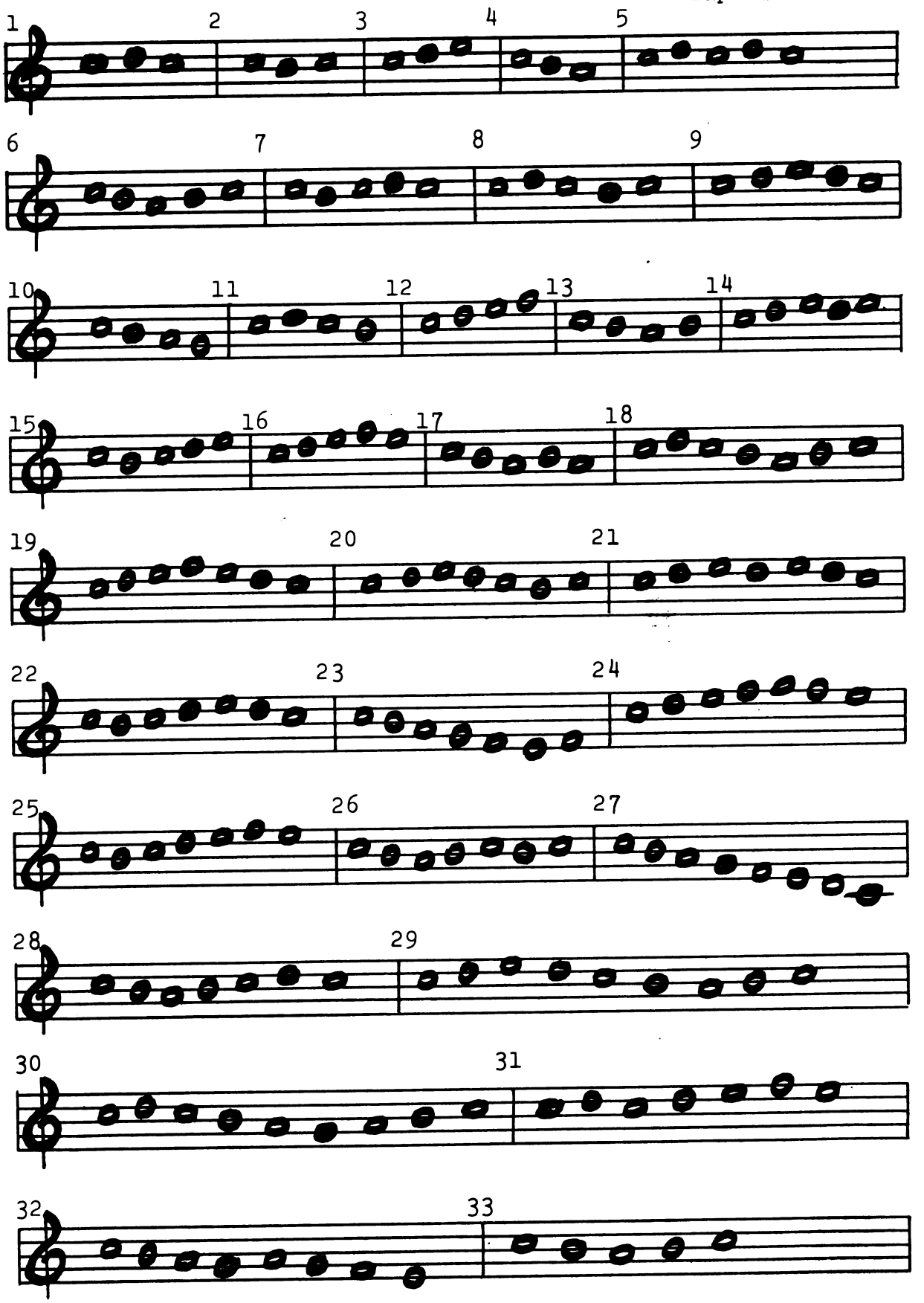
Reeves, William. "An Exploratory Study of Two Sets of Theories of Guthrie and Wheeler as they Relate to the Development of Instrumental Musicianship." Unpublished Ph.D. dissertation, University of California, 1954.

Spohn, Charles L. "An Exploration in the Use of Recorded Teaching Material to Develop Aural Comprehension in College Music Classes." Unpublished Ph.D. dissertation, Ohio State University, Columbus, Ohio, 1959.

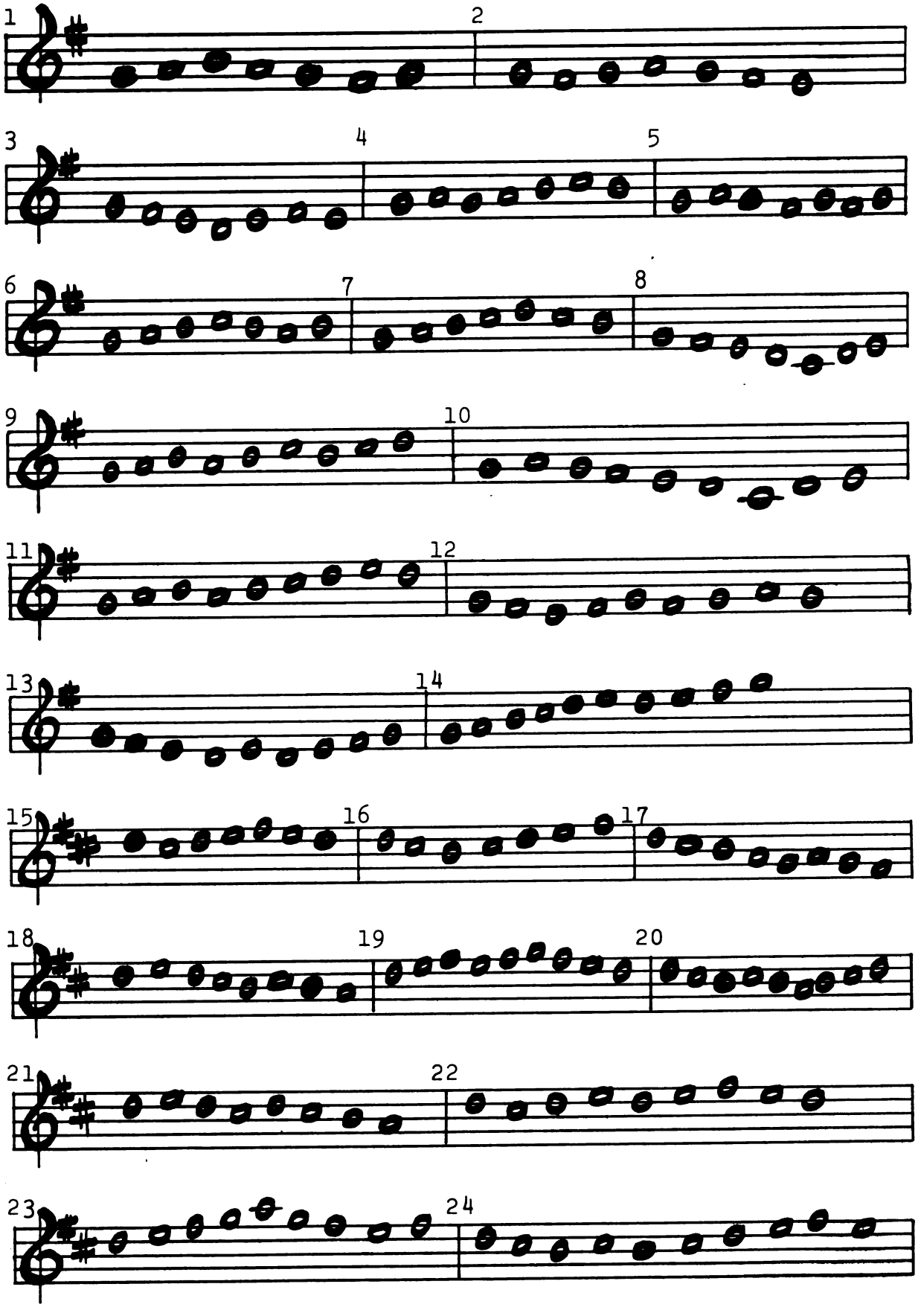
APPENDIX A

PRACTICE MATERIALS OF THE STUDY

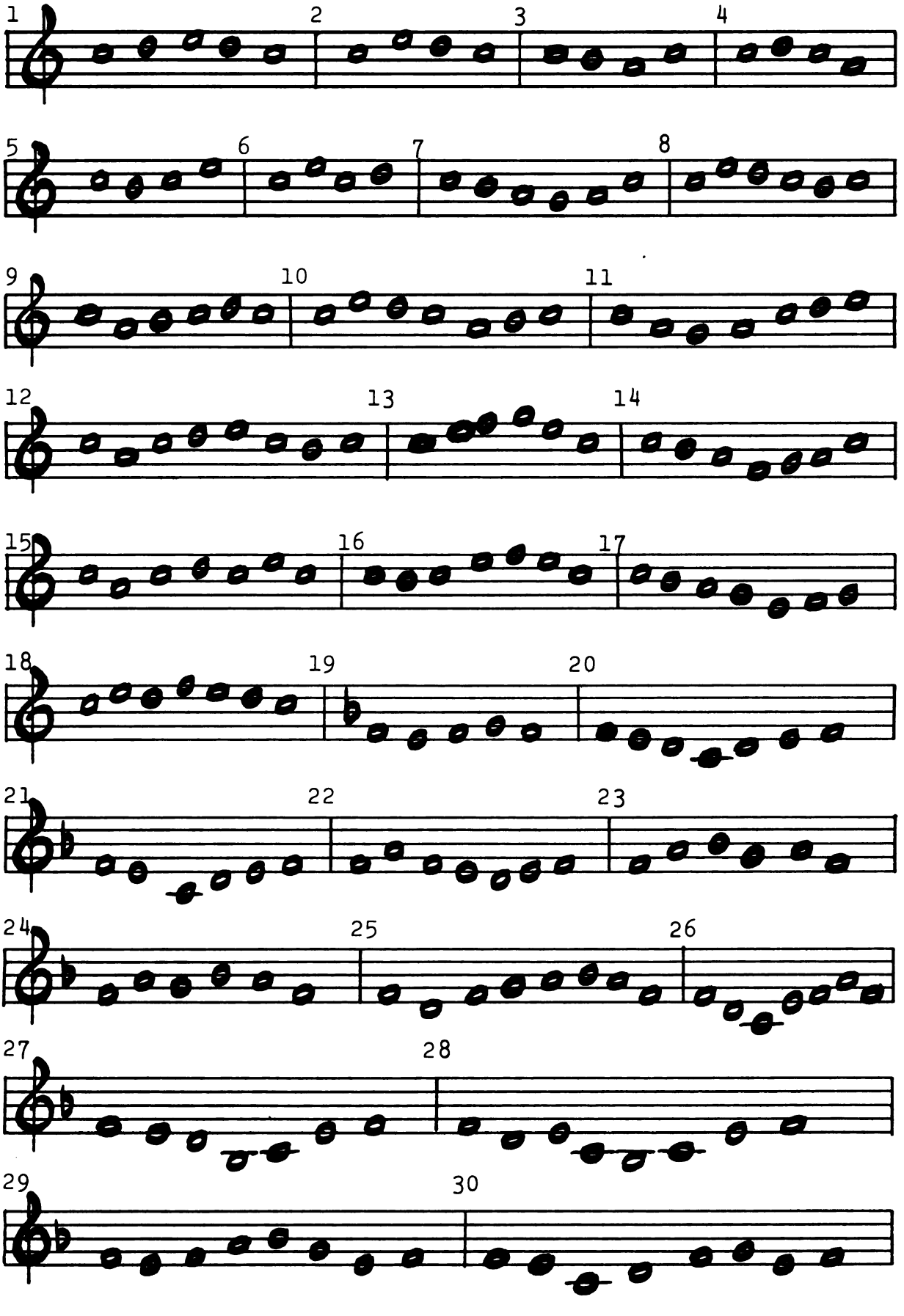
Series A
Tape #1



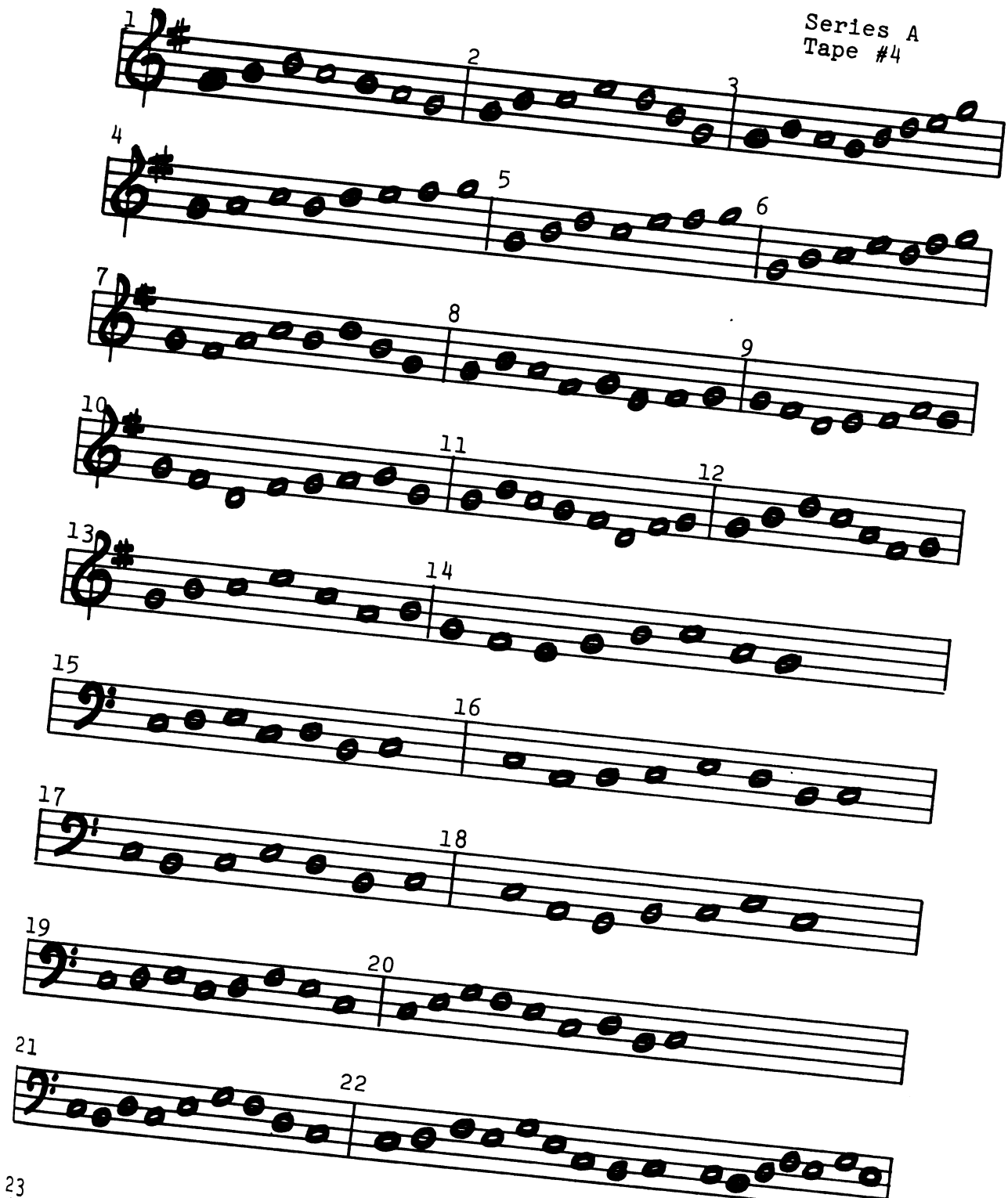
Series A
Tape #2



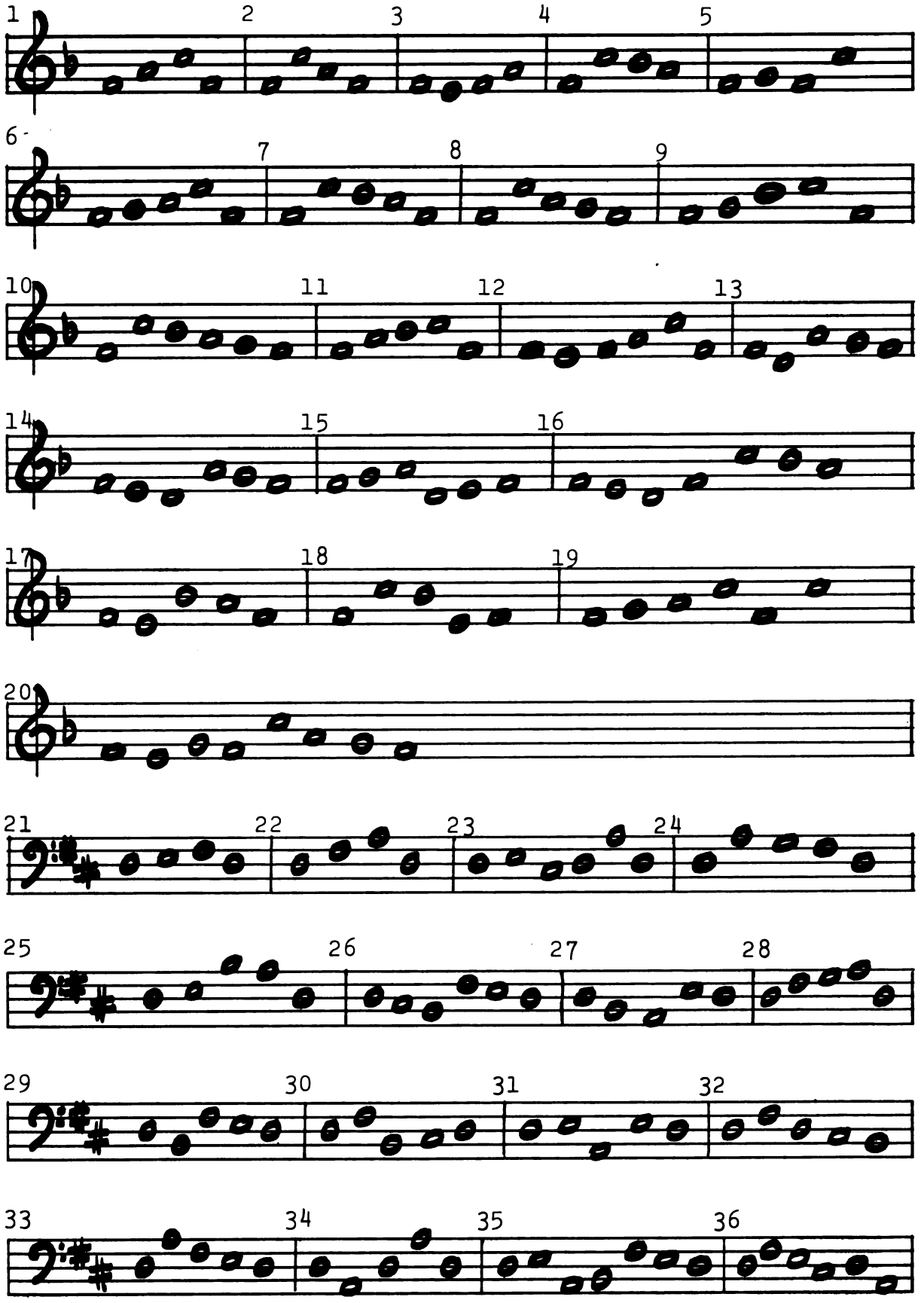
Series A
Tape #3



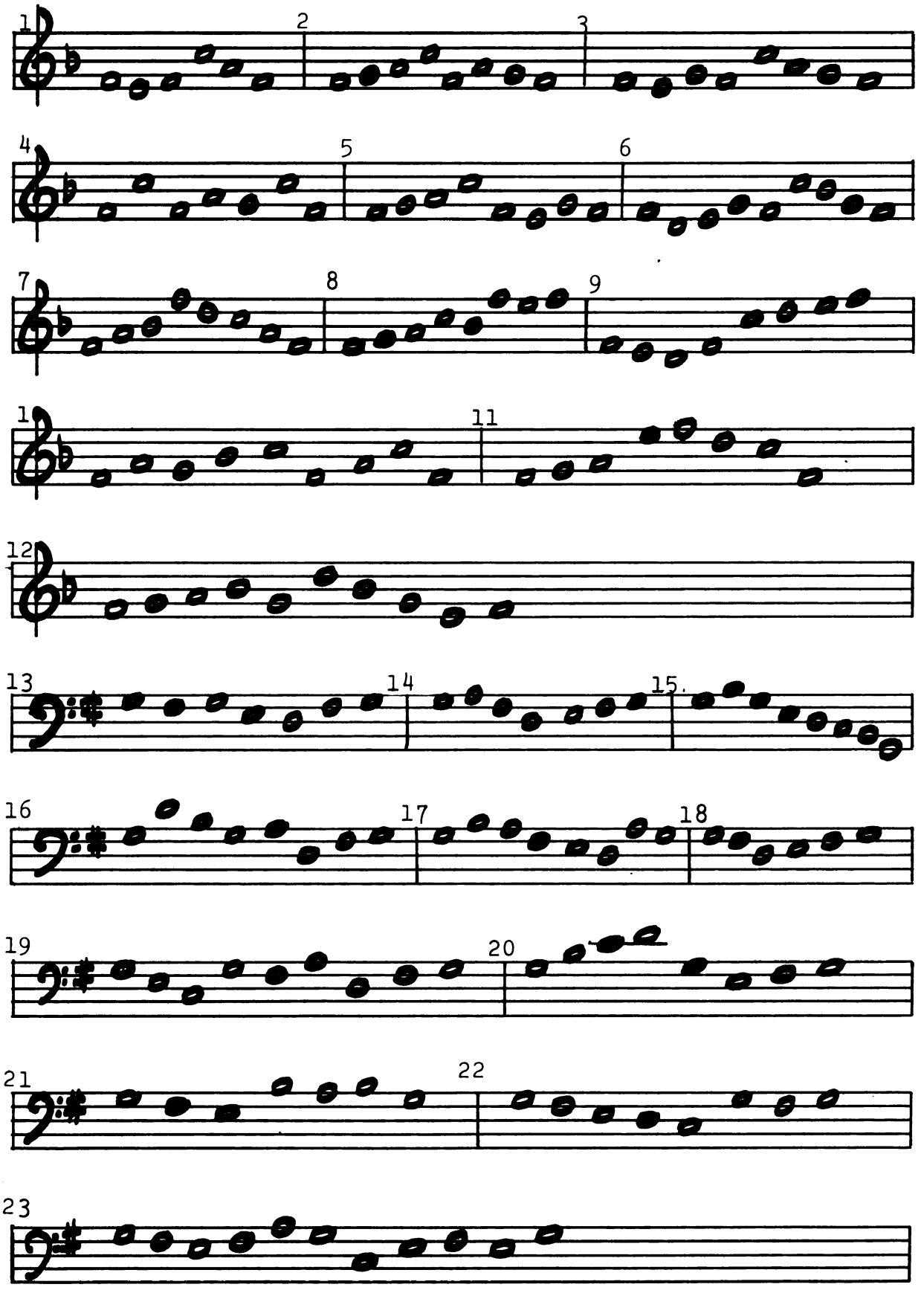
Series A
Tape #4



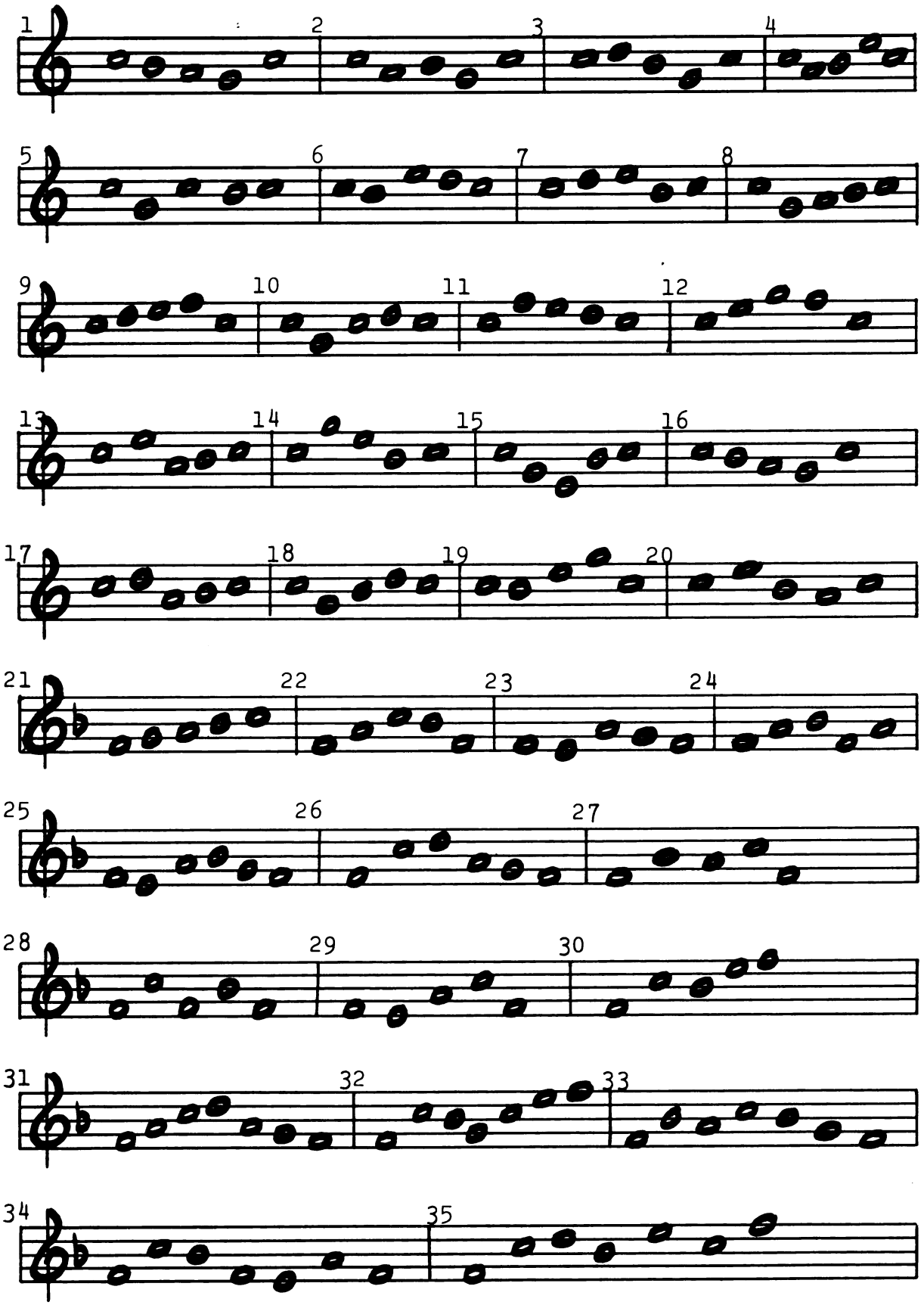
Series A
Tape #5



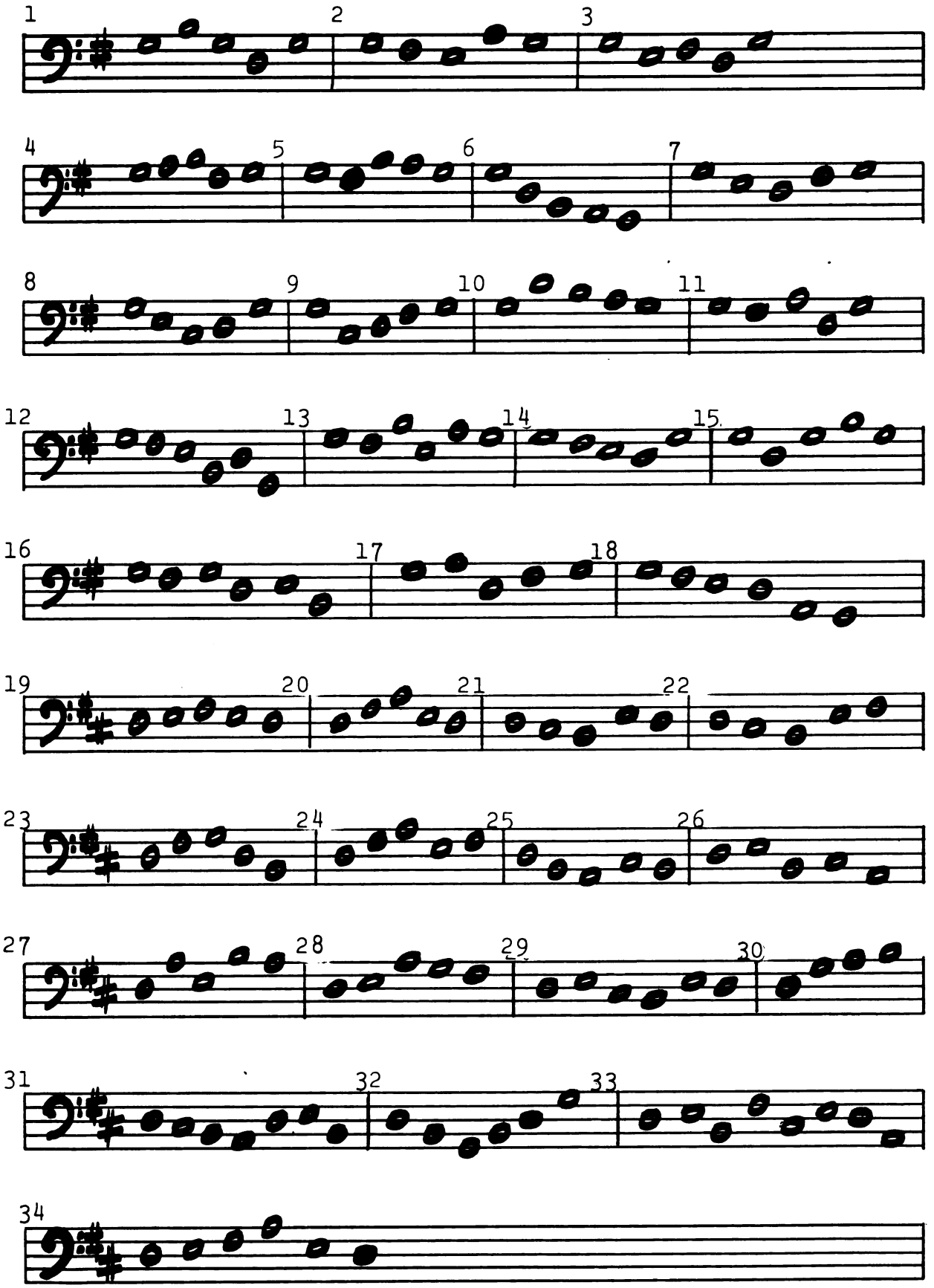
Series A
Tape #6



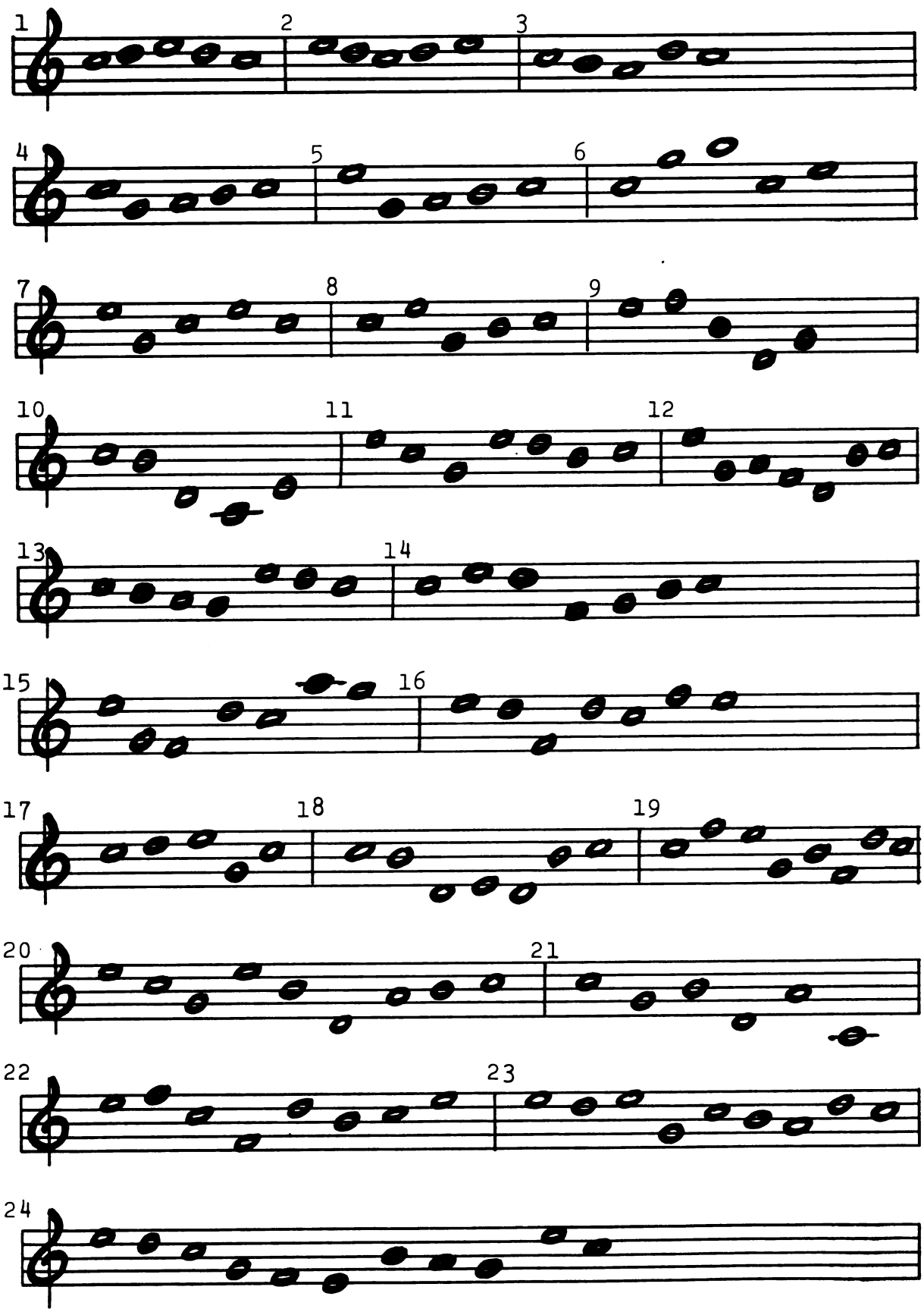
Series A
Tape #7



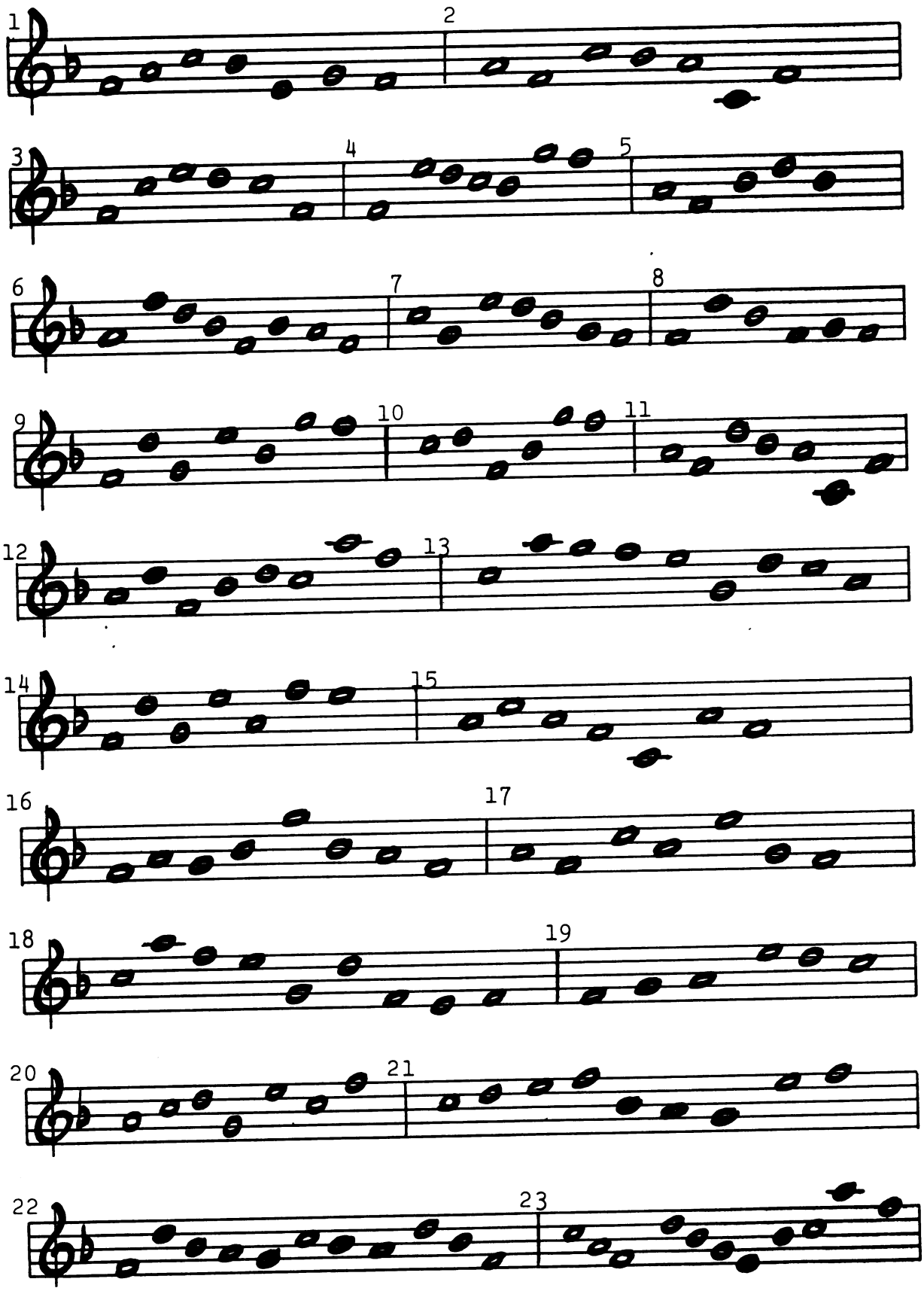
Series A
Tape #8



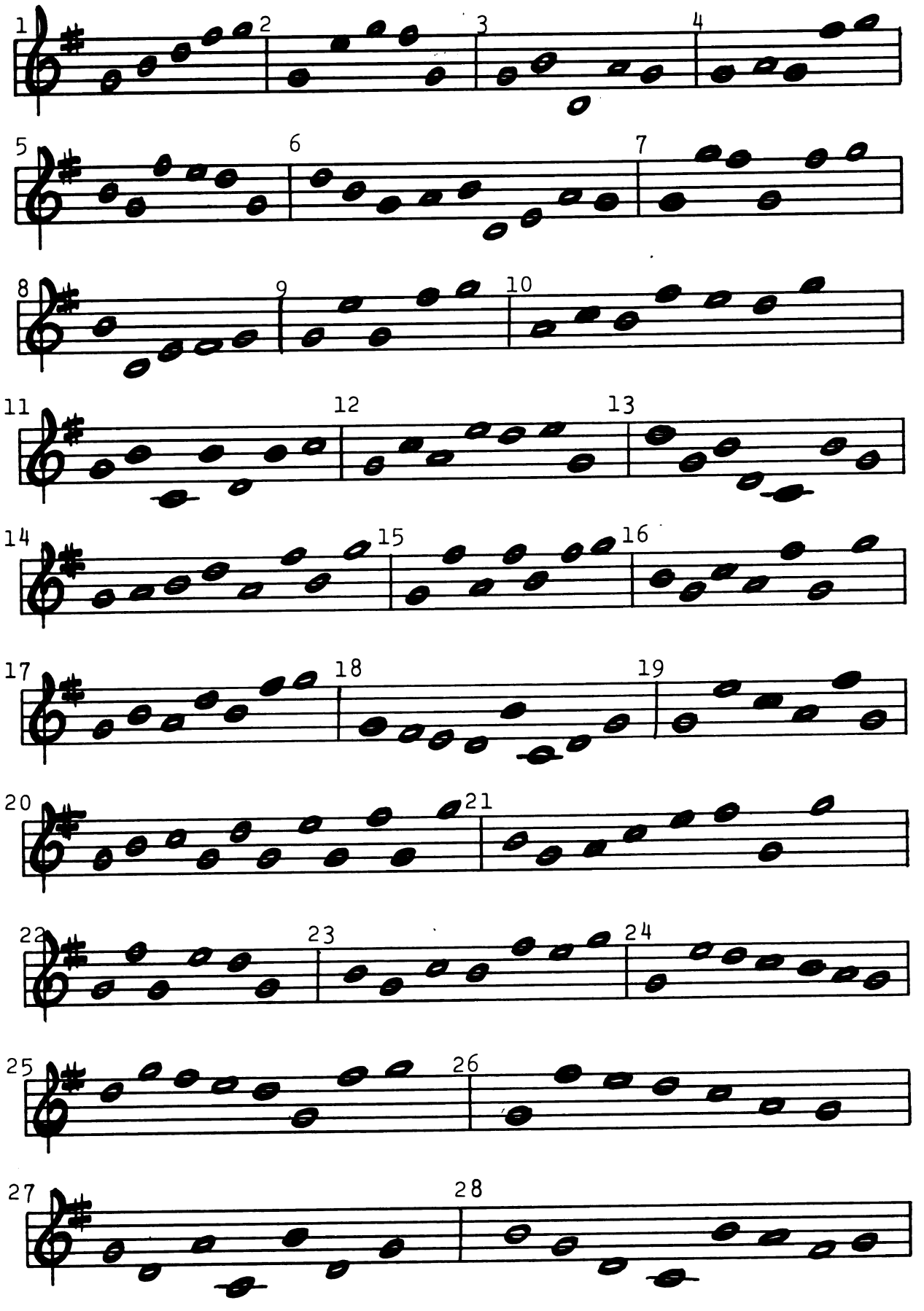
Series A
Tape #9



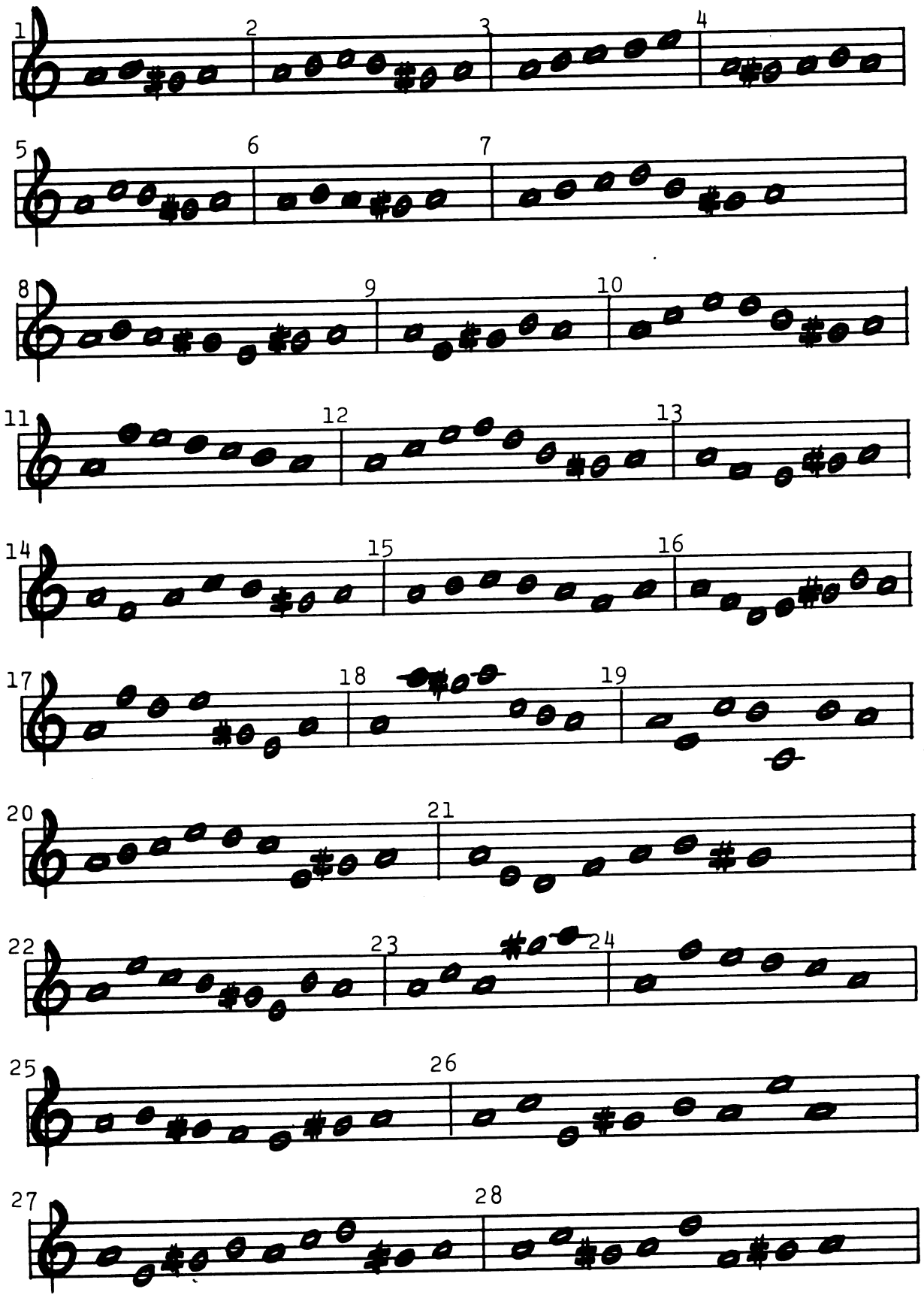
Series A
Tape #10



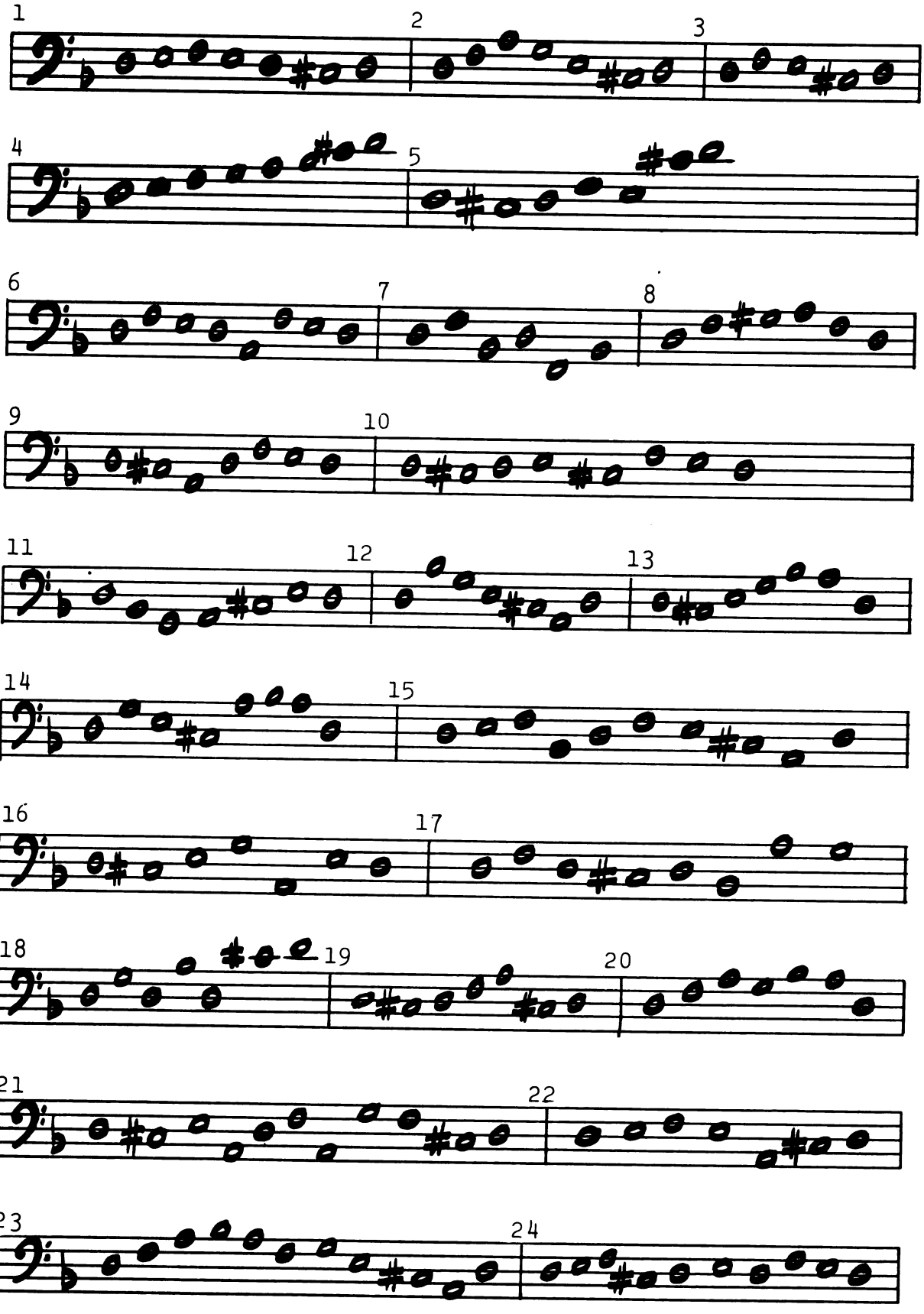
Series A
Tape #11



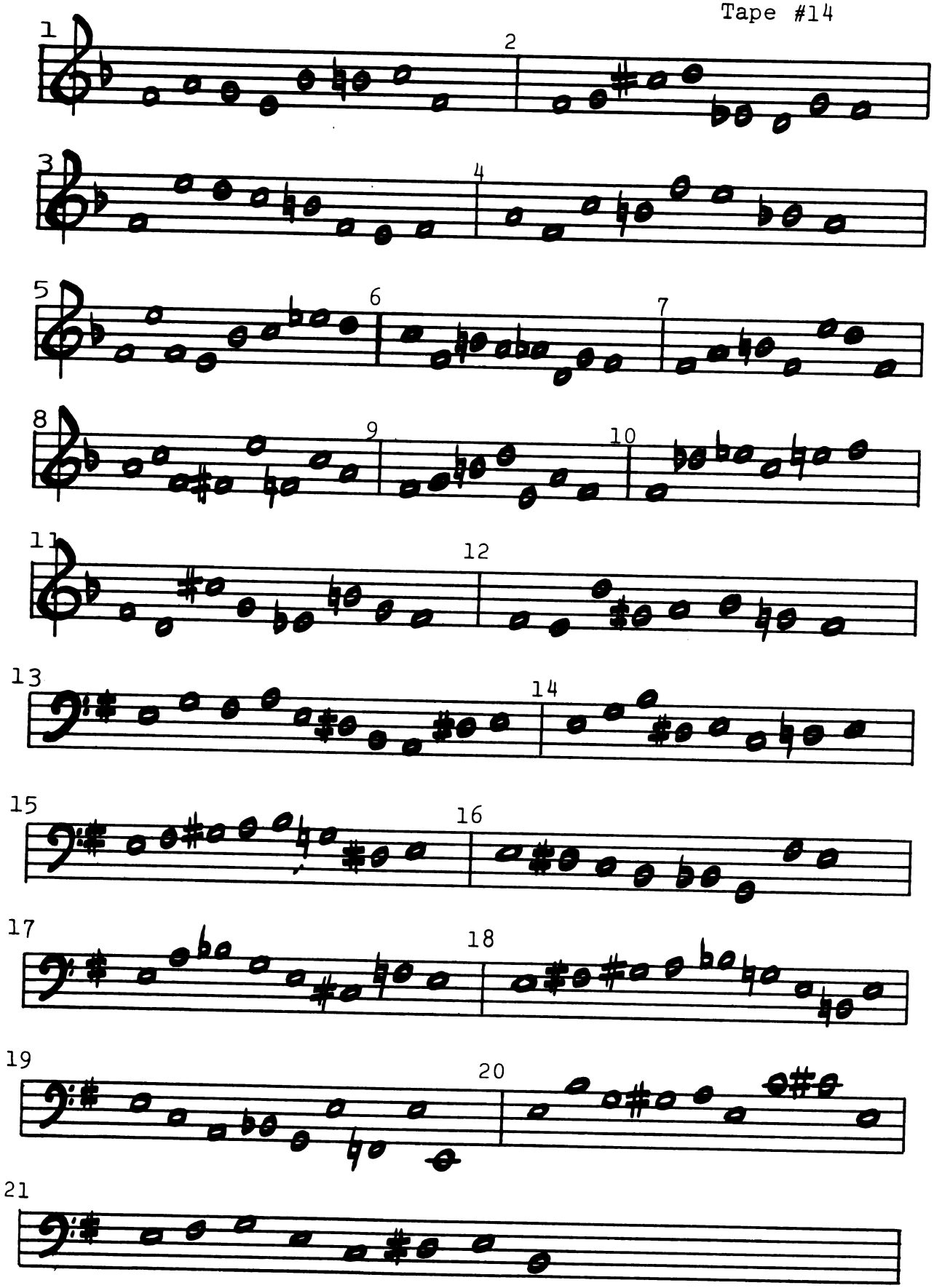
Series A
Tape #12



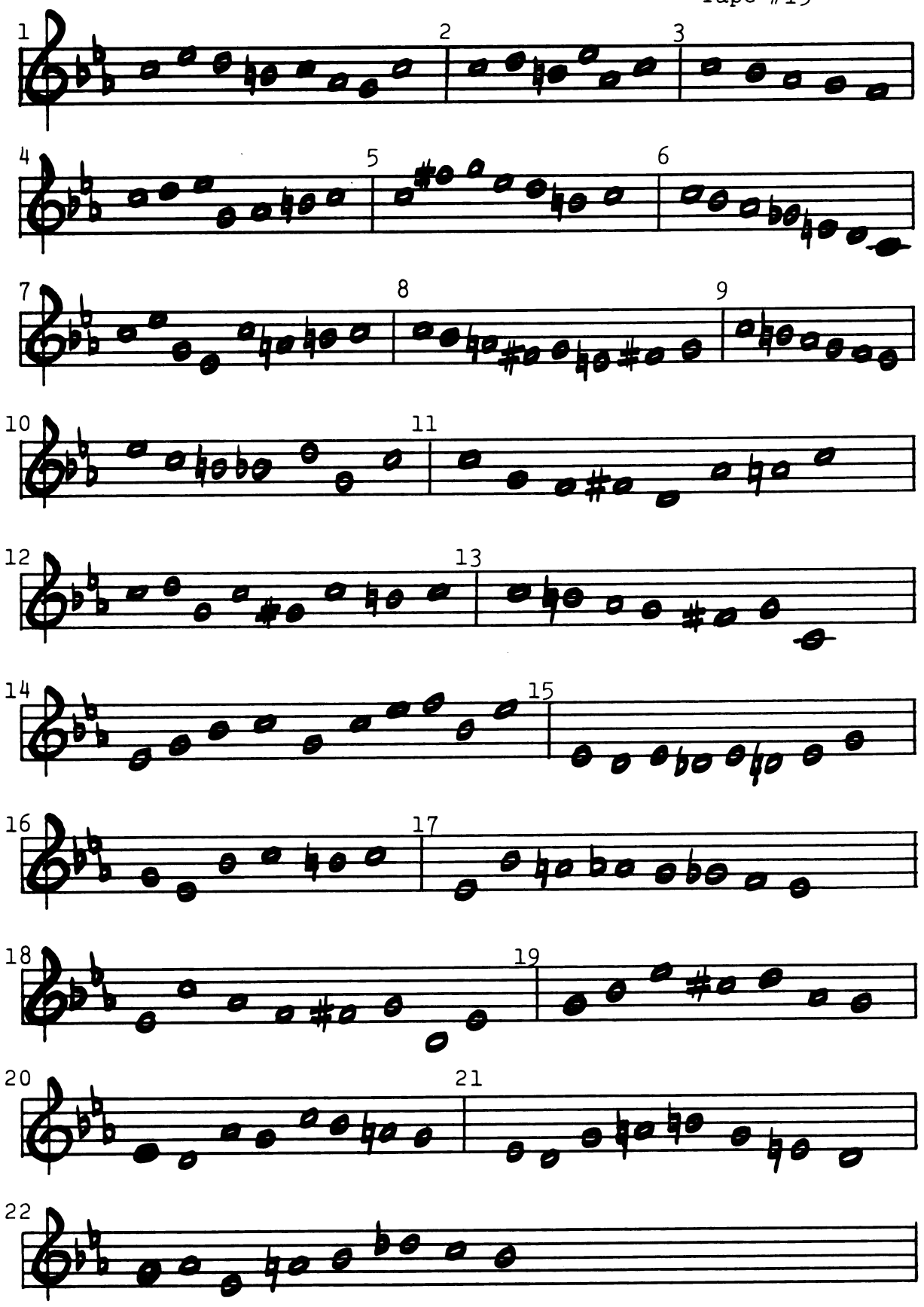
Series A
Tape #13



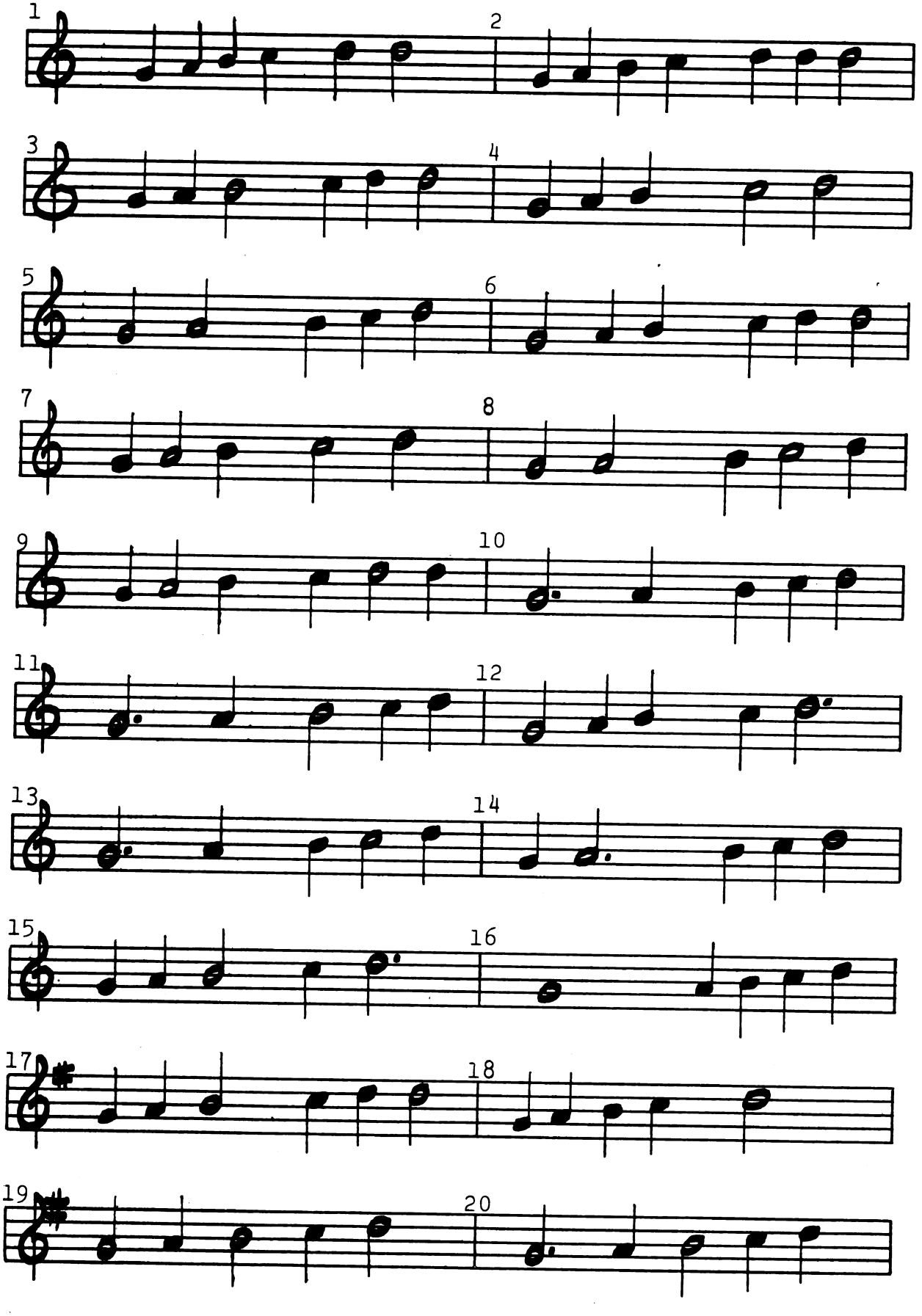
Series A
Tape #14



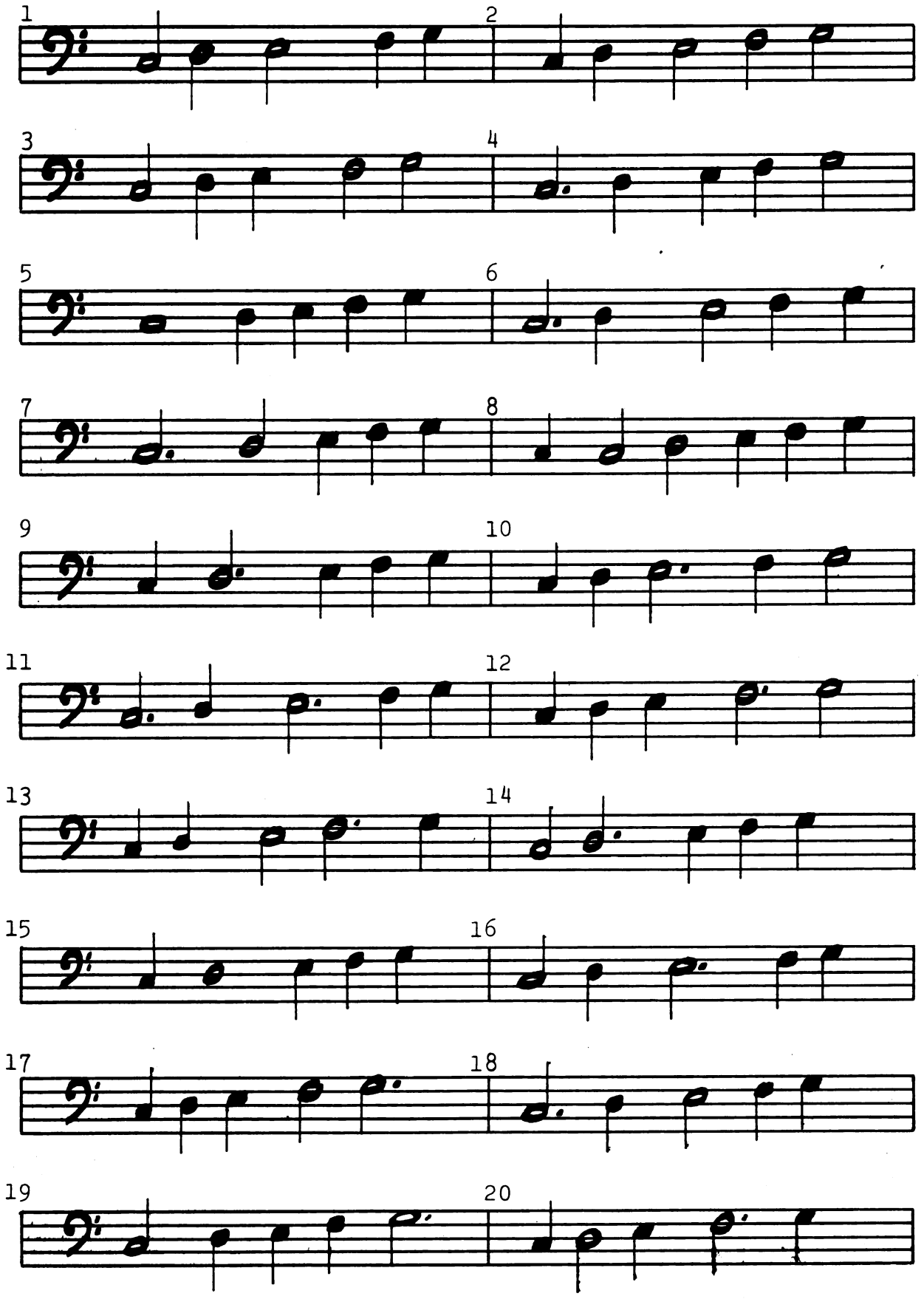
Series A
Tape #15



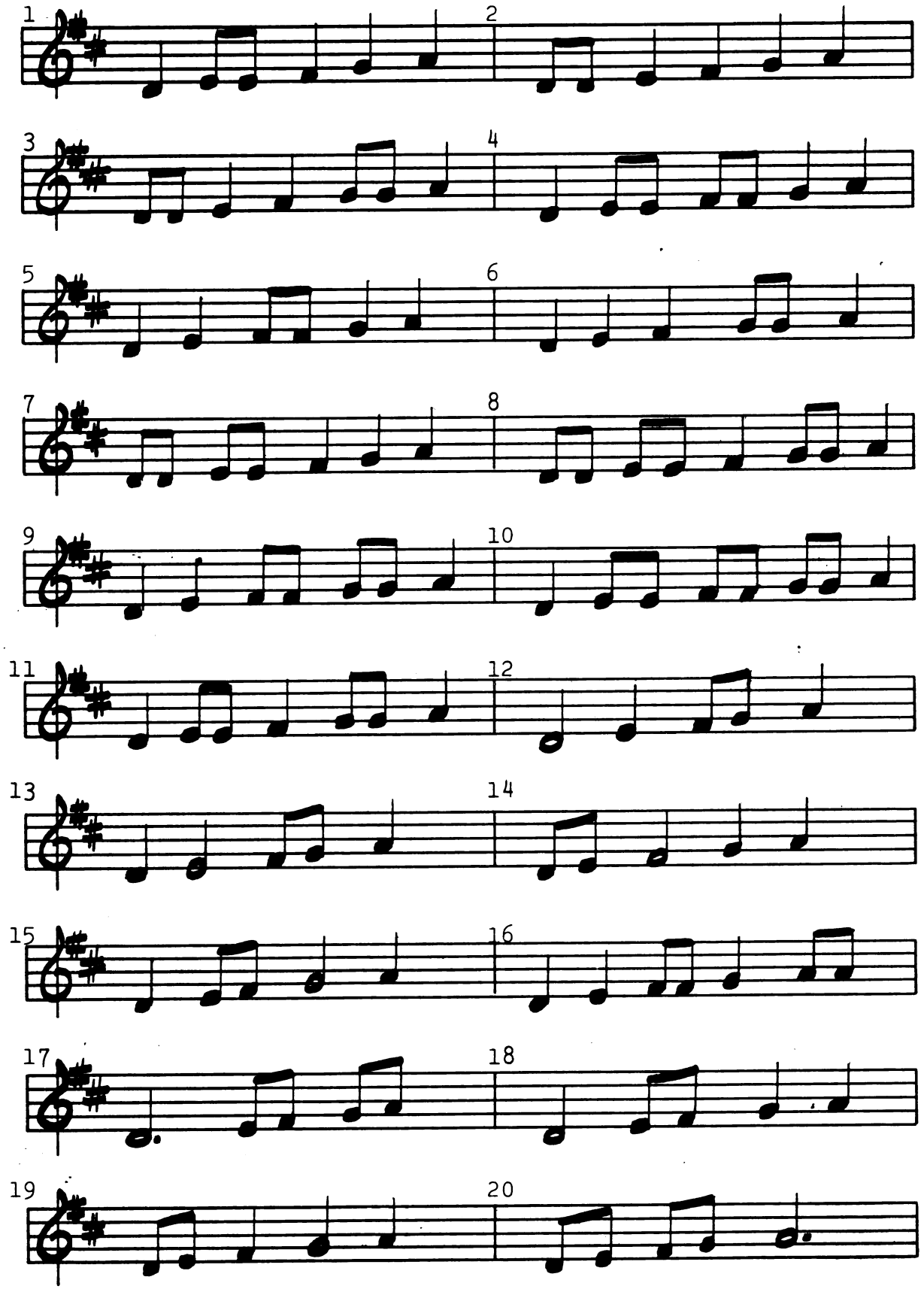
Series B
Tape #1



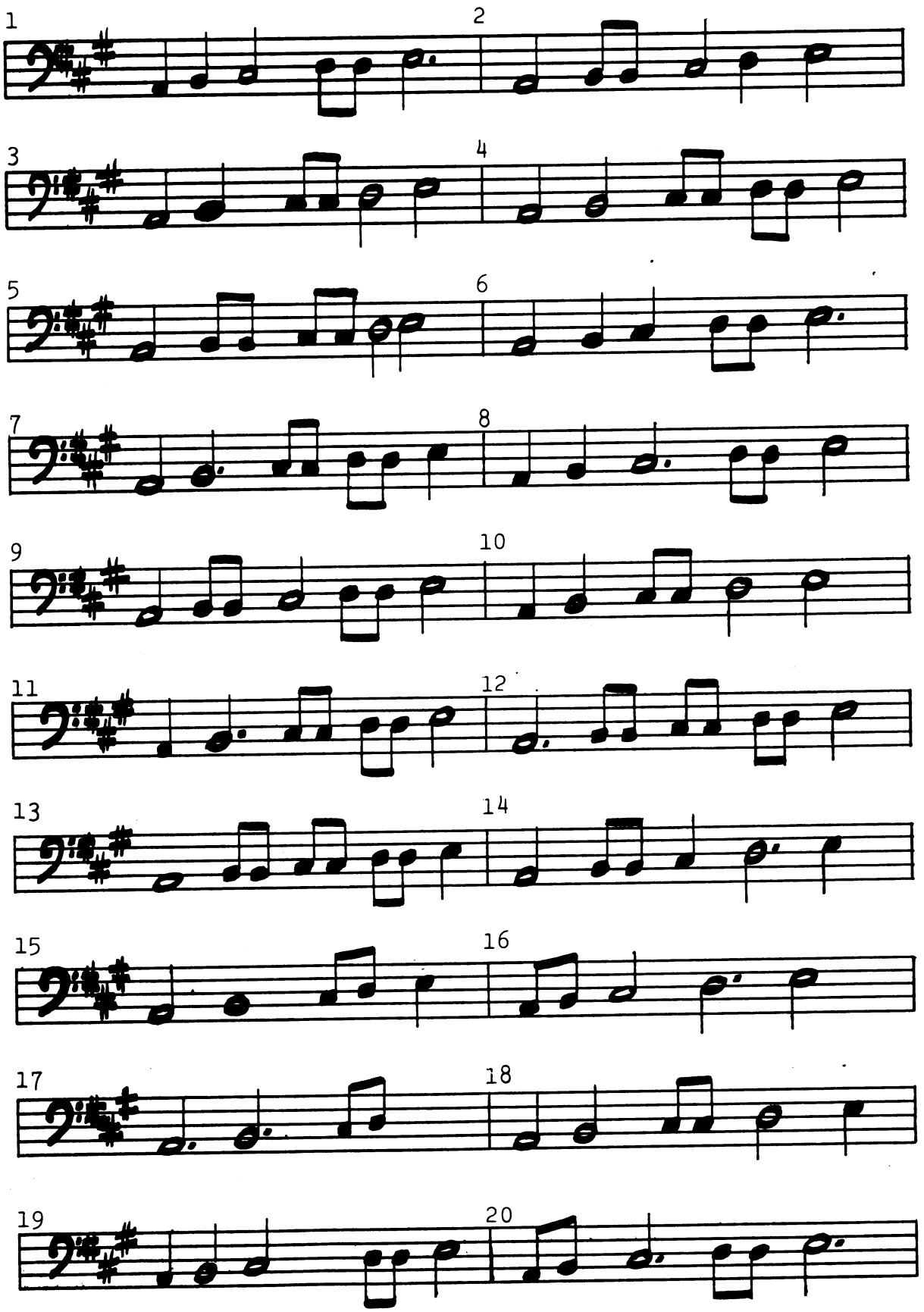
Series B
Tape #2



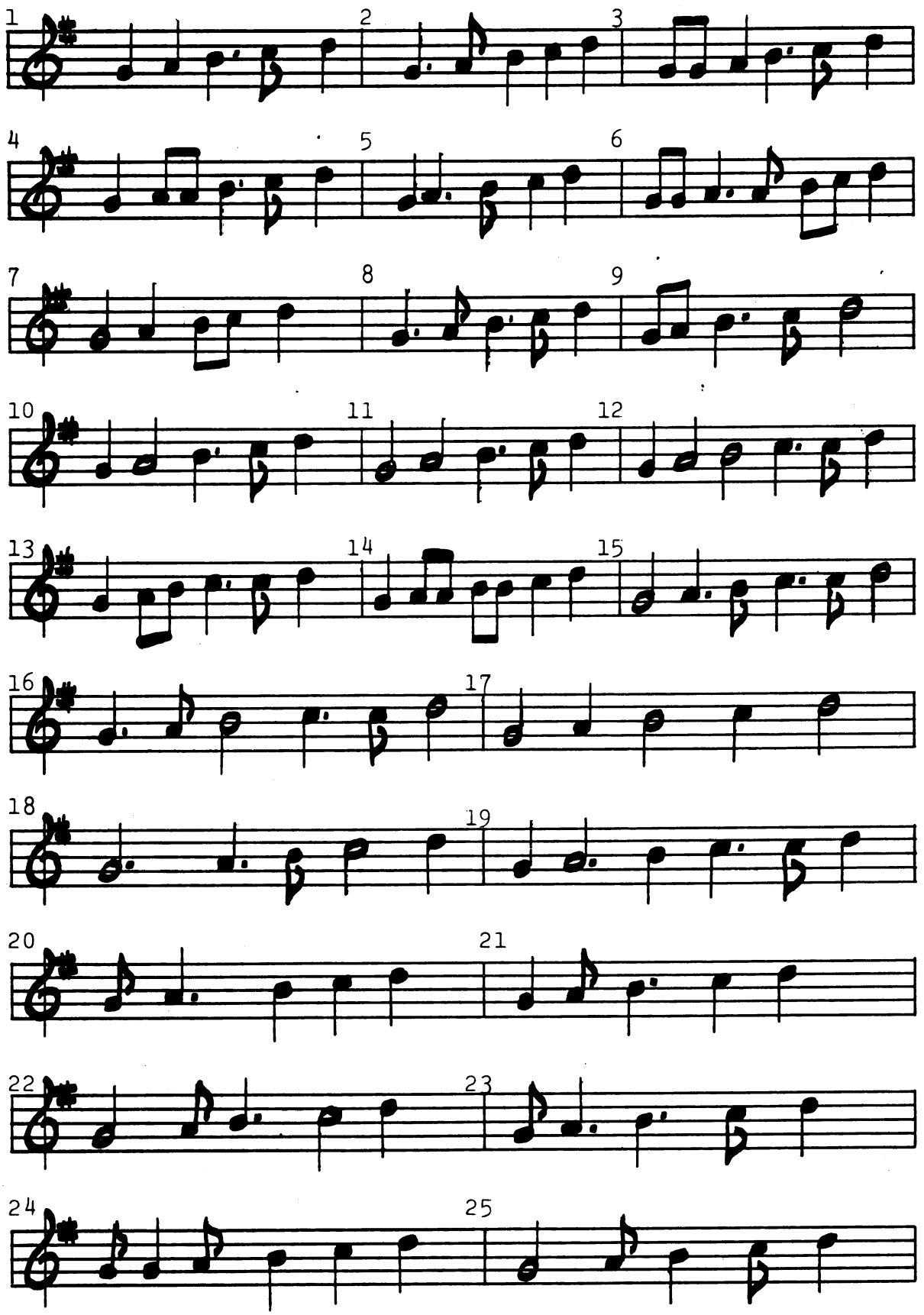
Series B
Tape #3



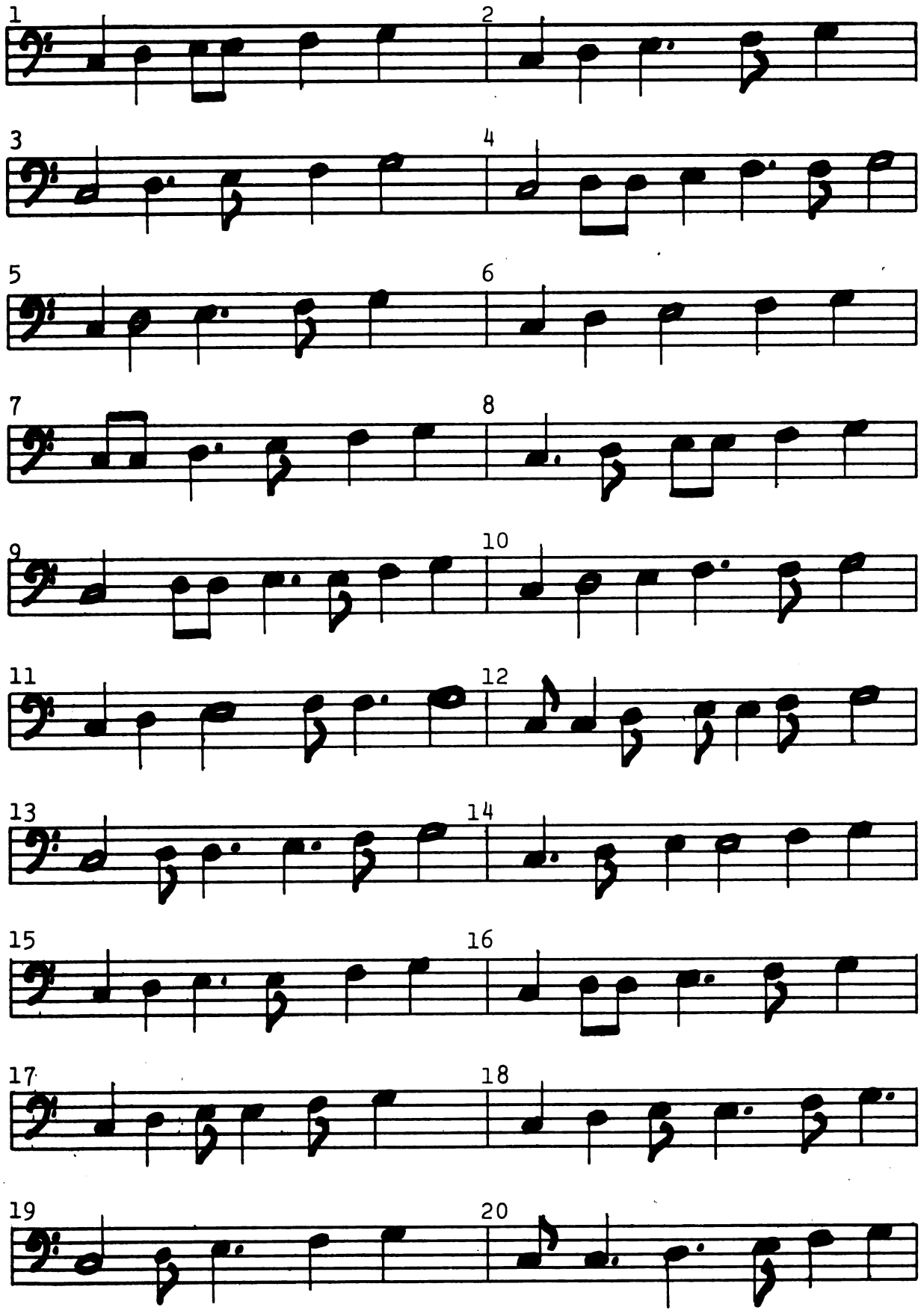
Series B
Tape #4



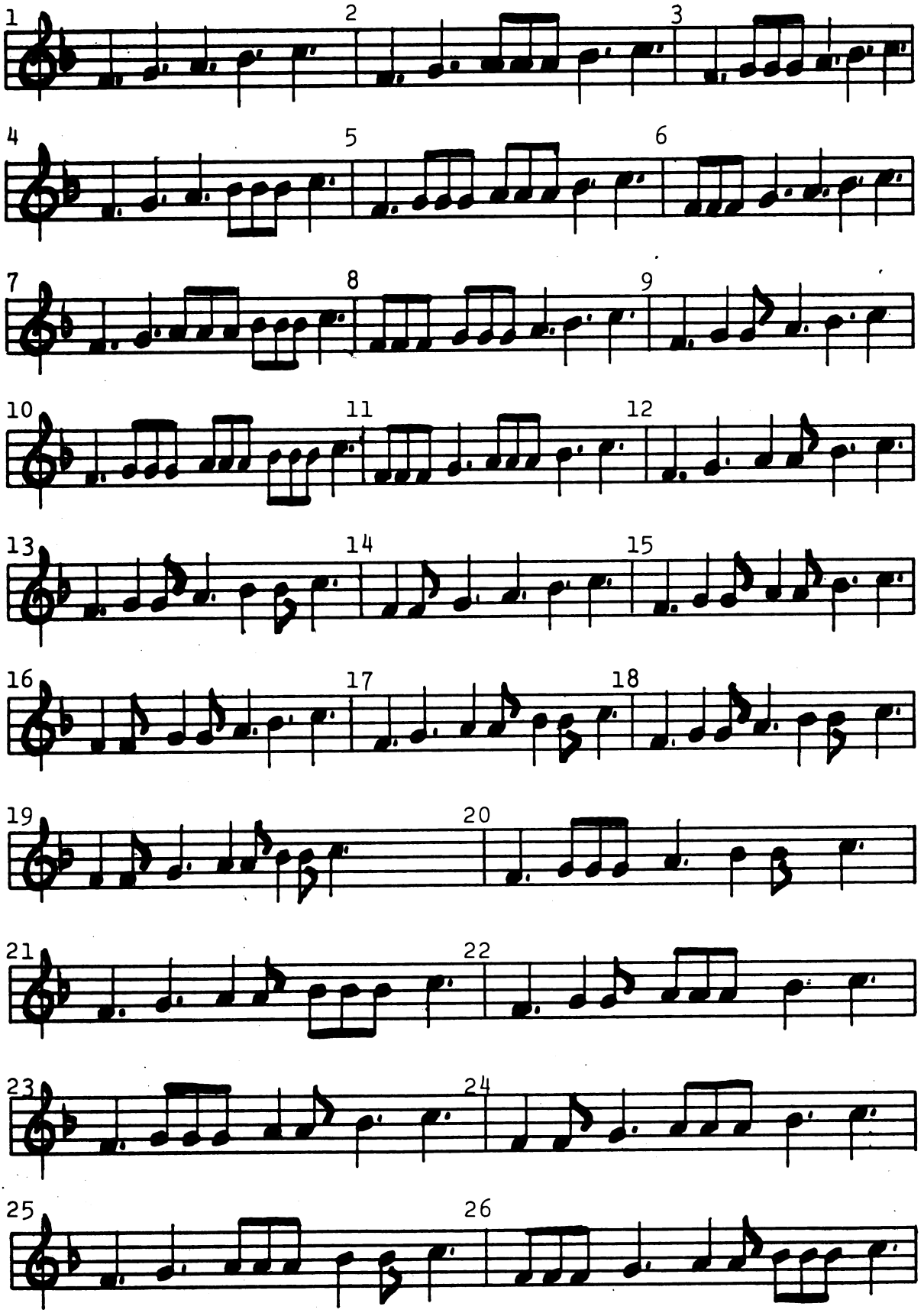
Series B
Tape #5



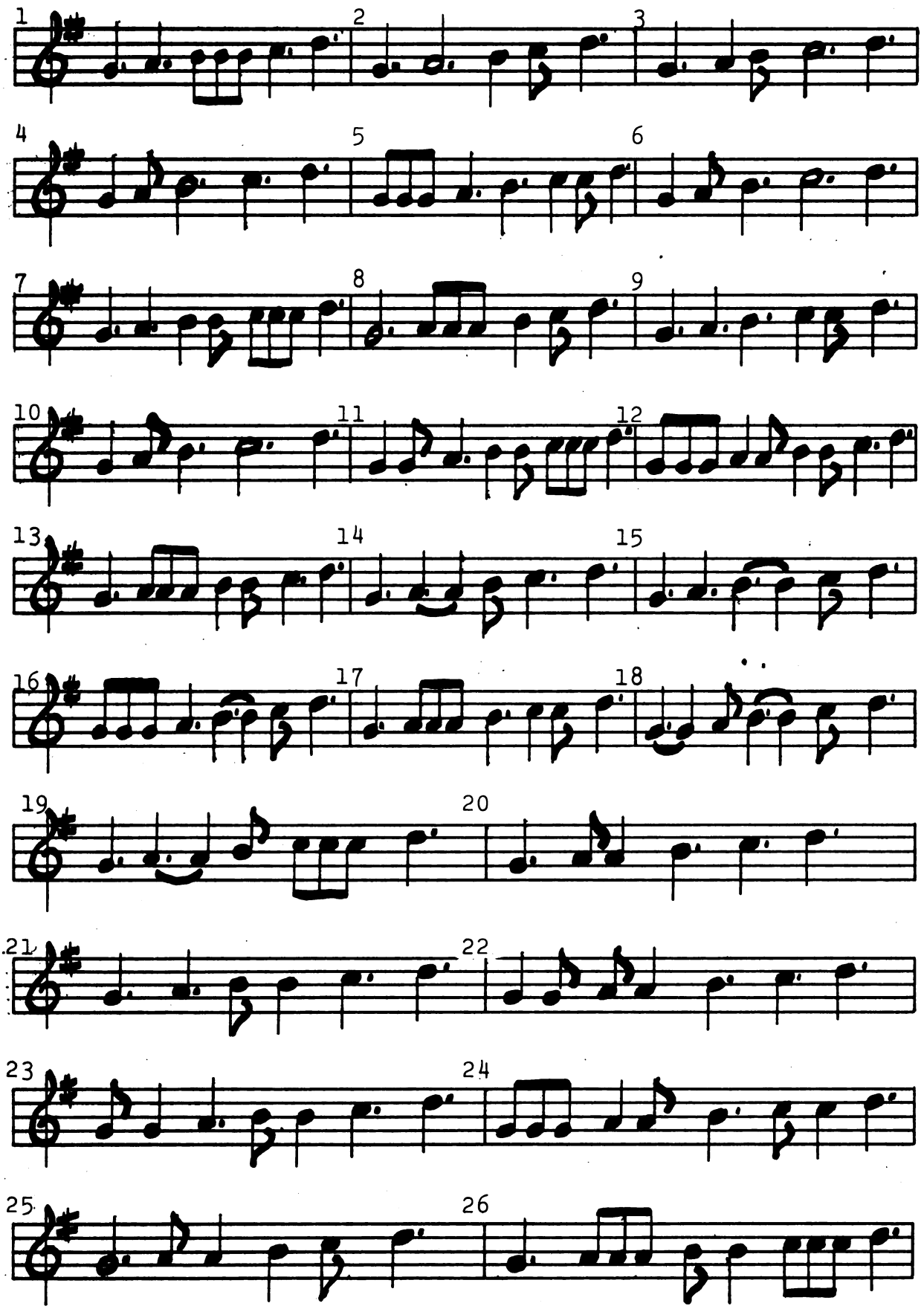
Series B
Tape #6



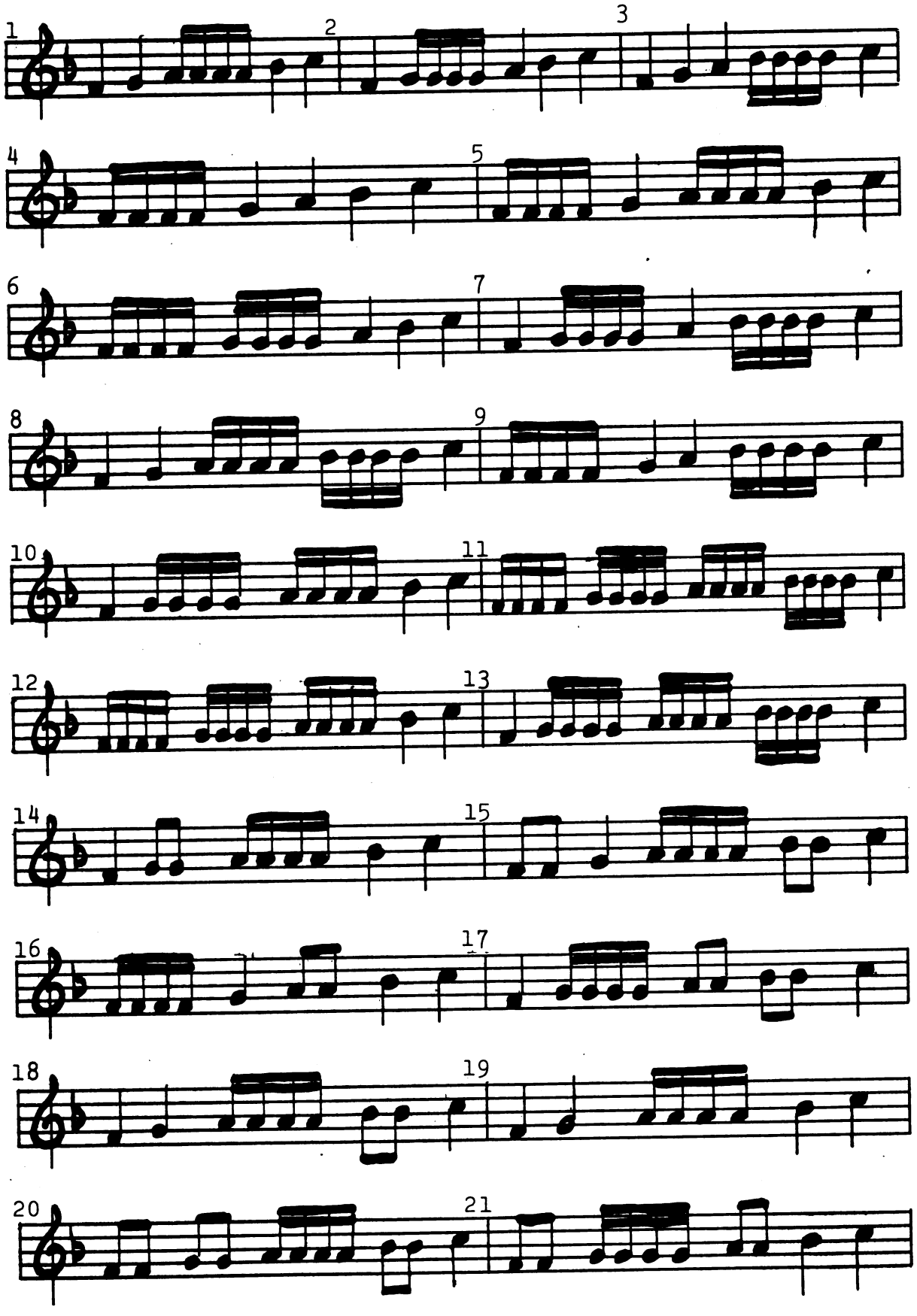
Series B
Tape #7



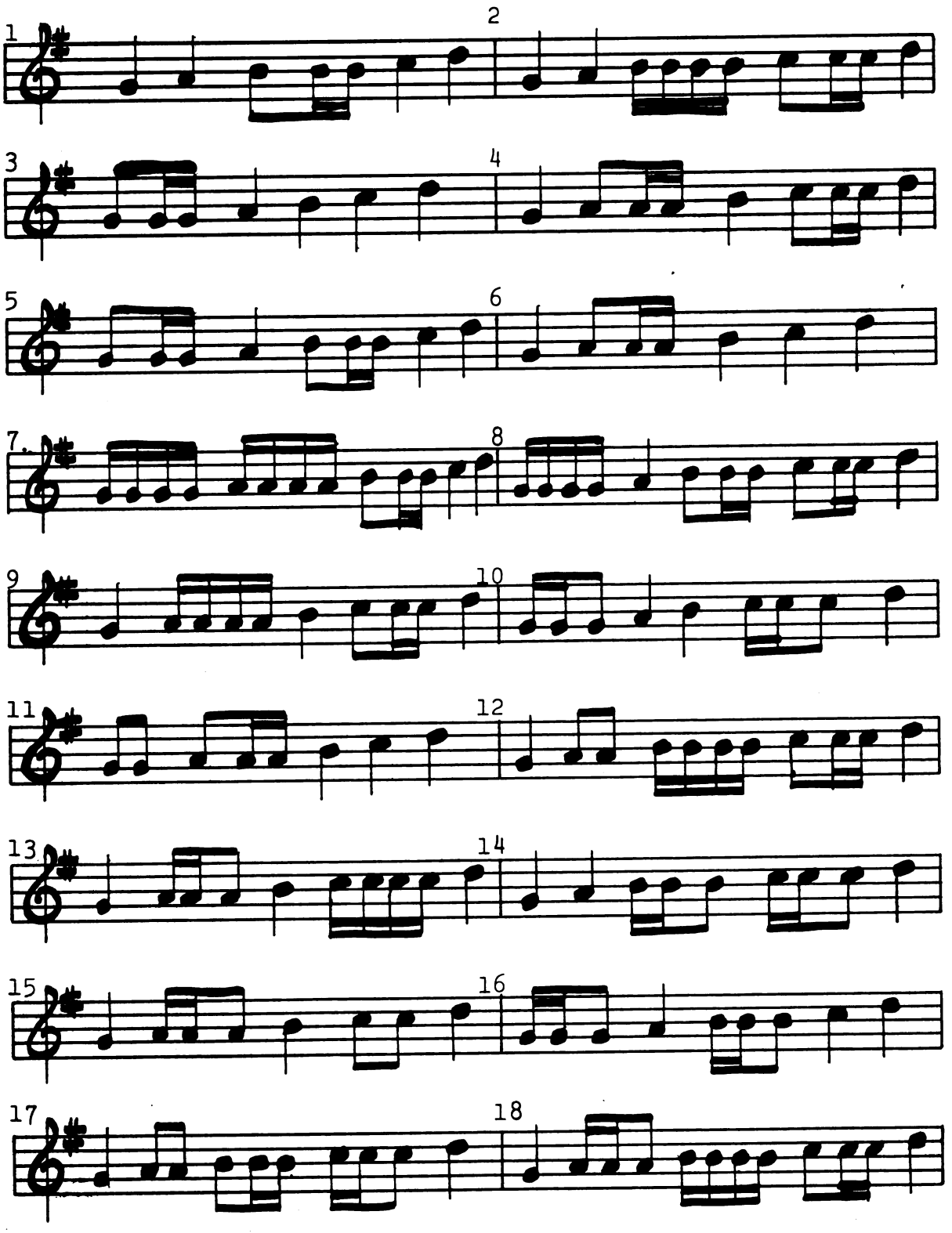
Series B
Tape #8



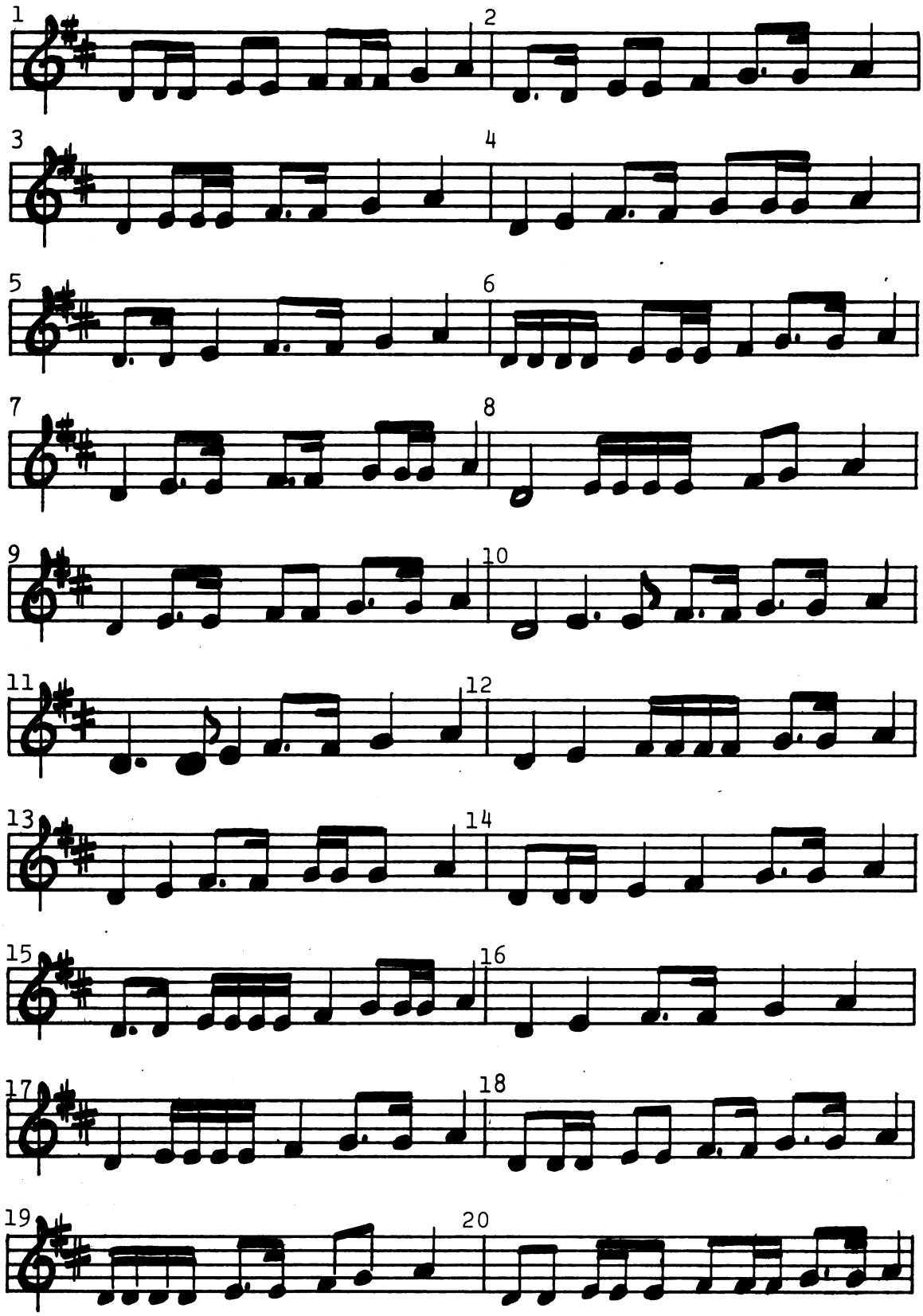
Series B
Tape #9



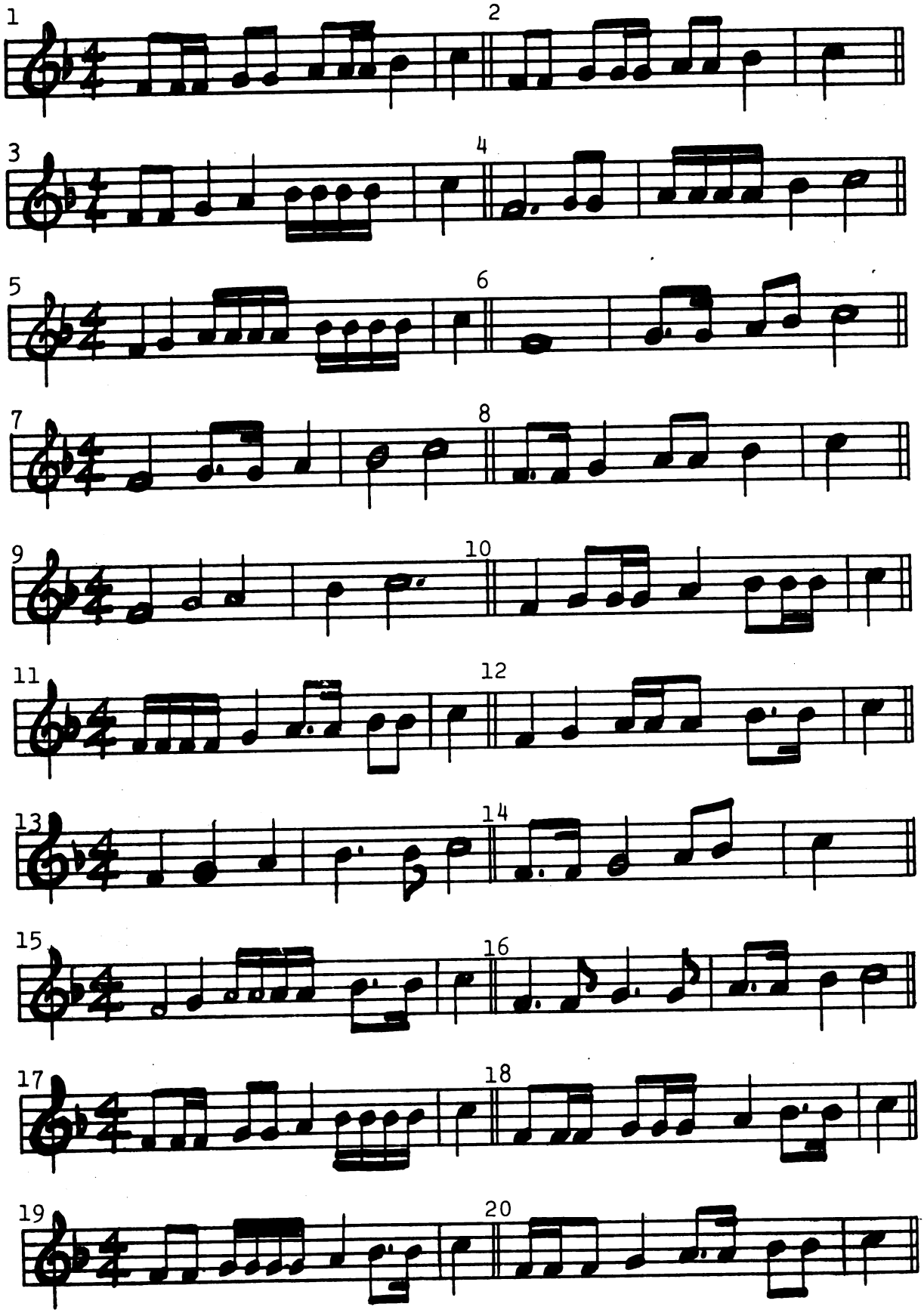
Series B
Tape #10



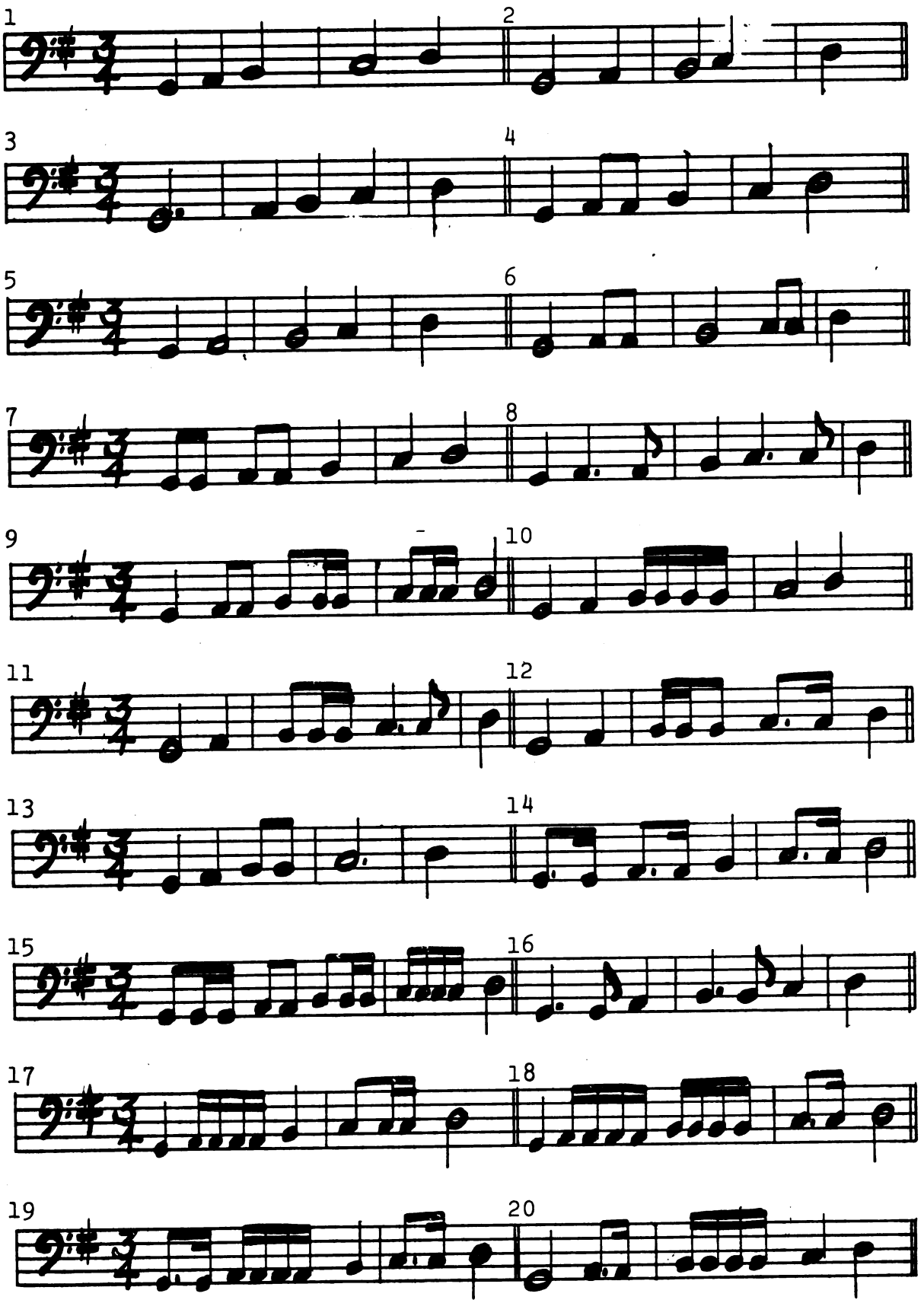
Series B
Tape #11



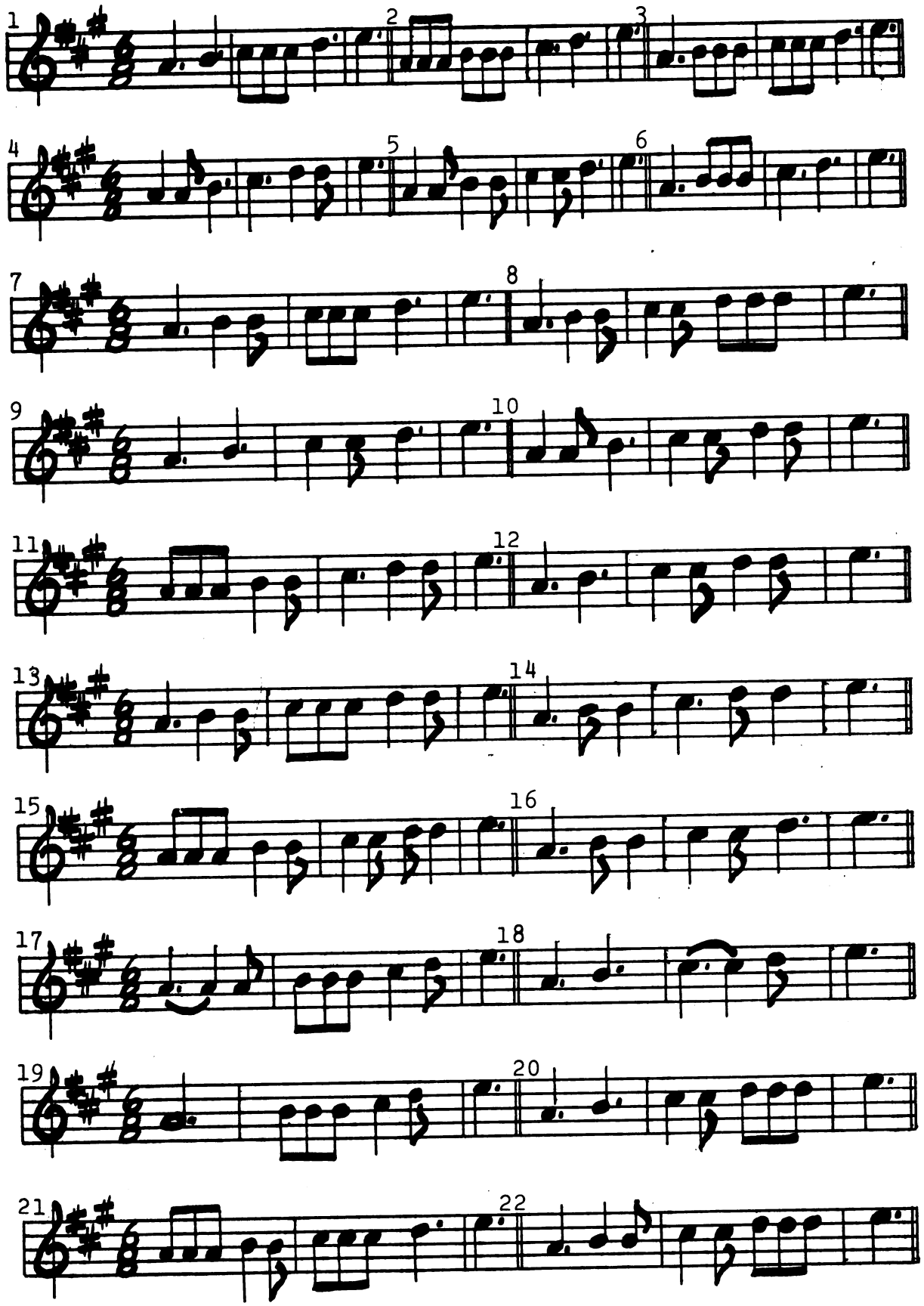
Series B
Tape #12



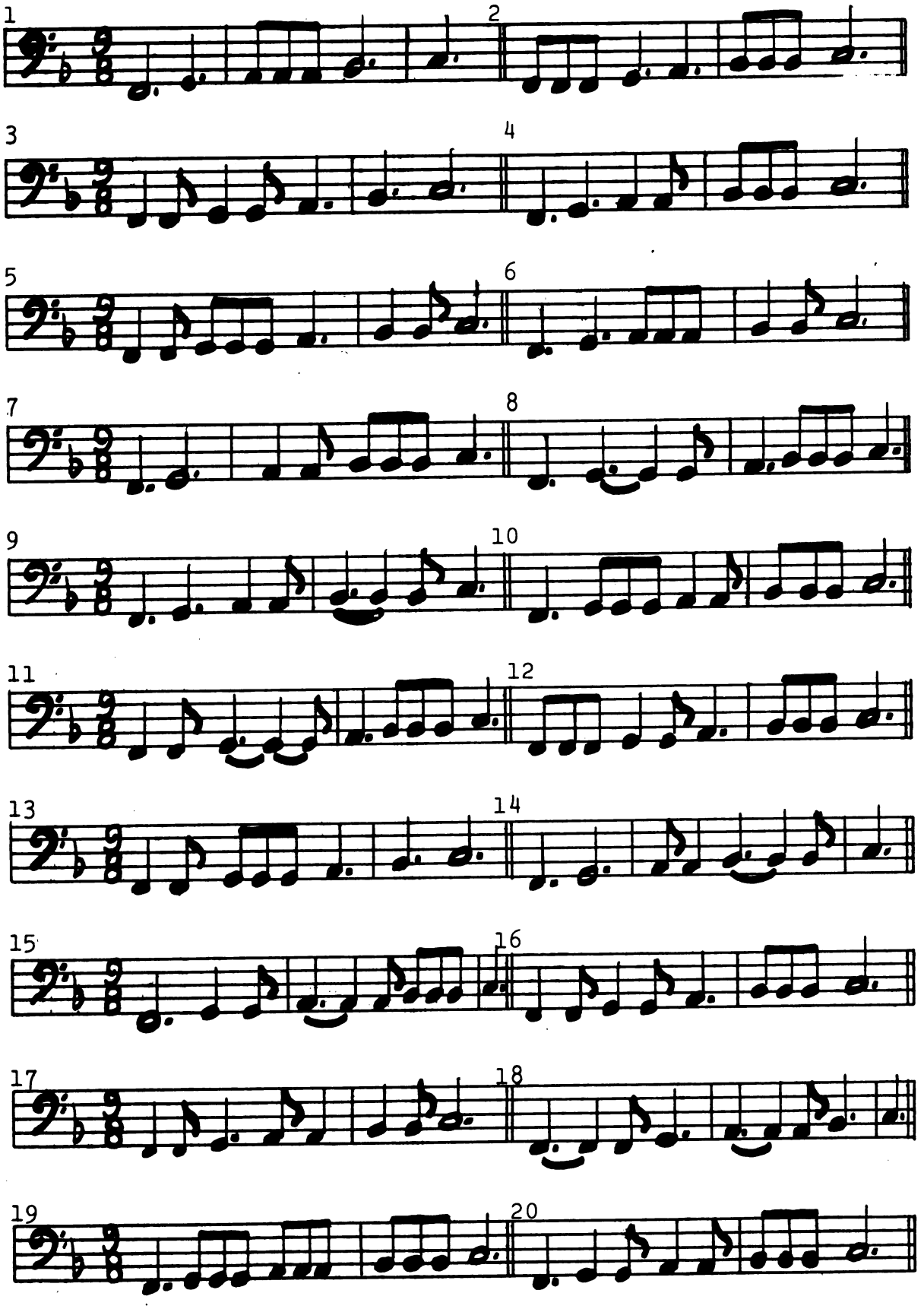
Series B
Tape #13



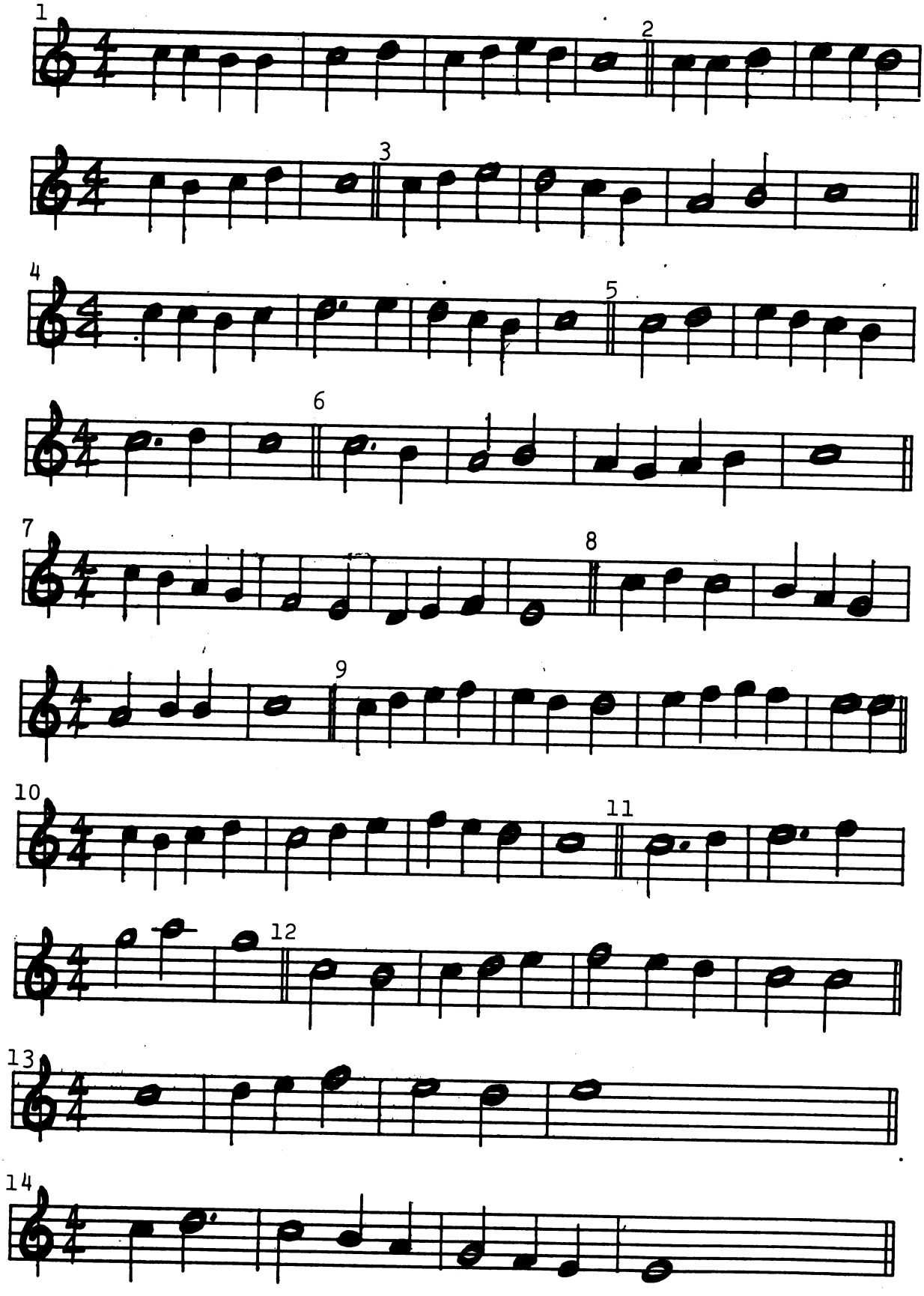
Series B
Tape #14



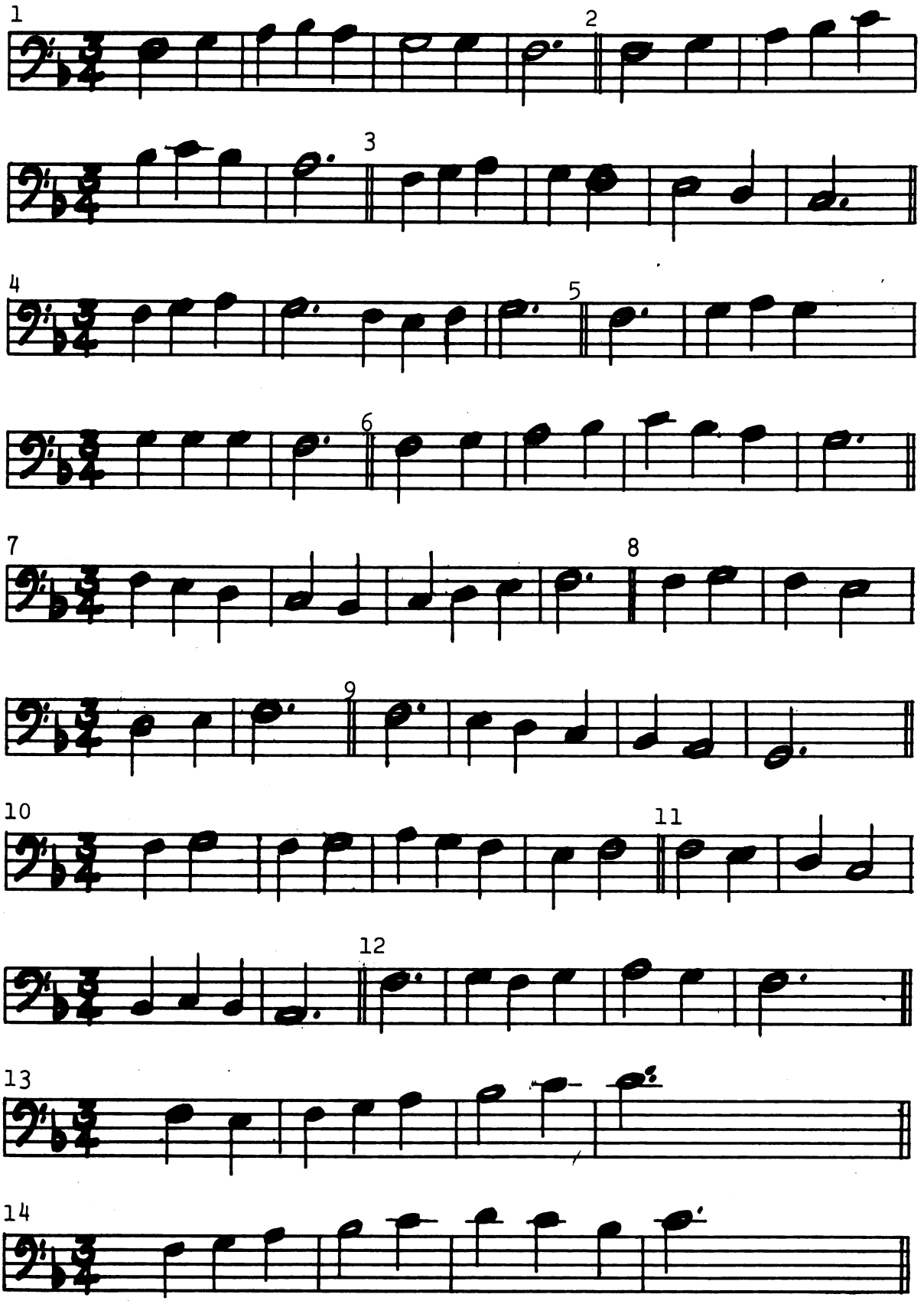
Series B
Tape #15



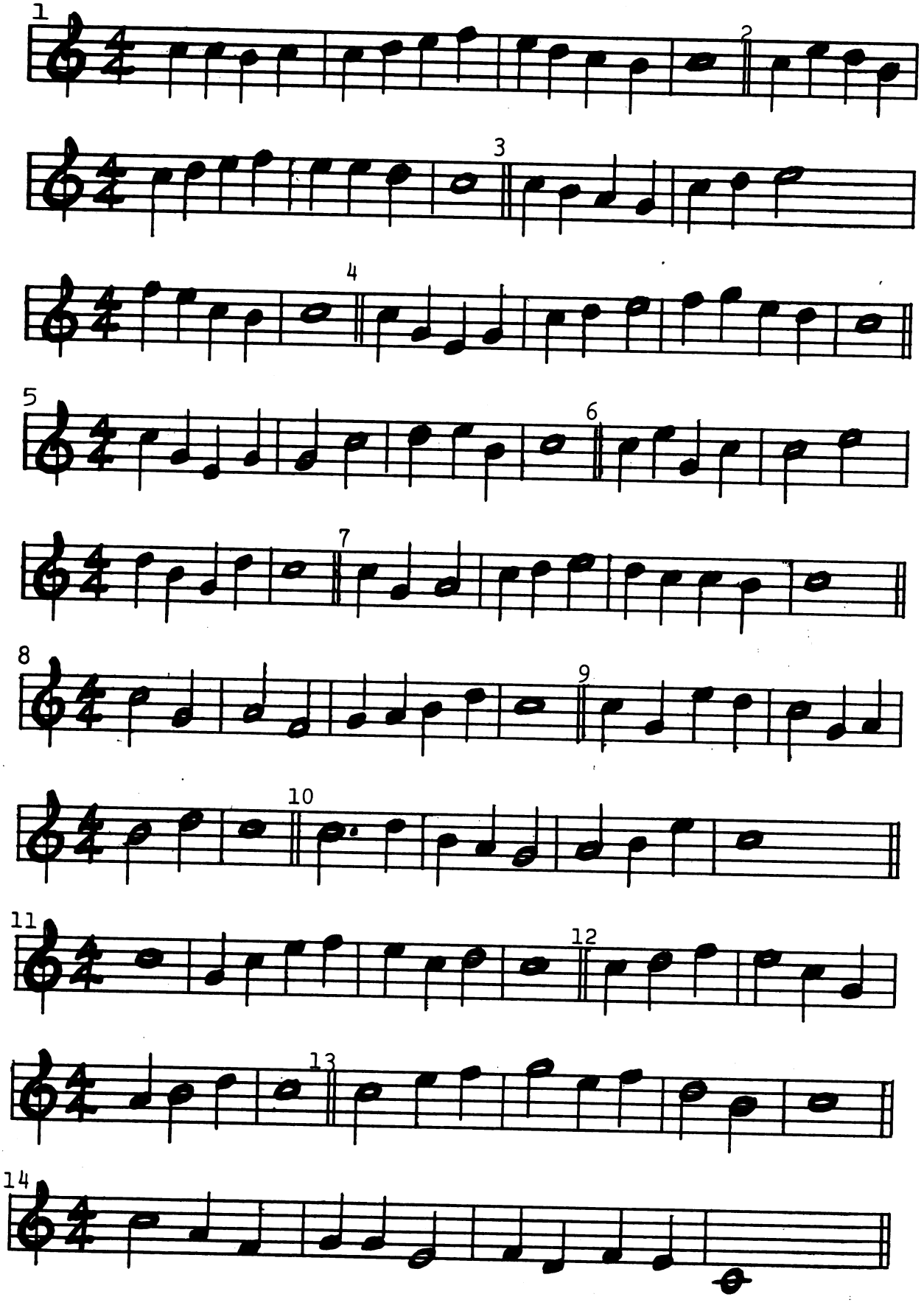
Series C
Tape #1



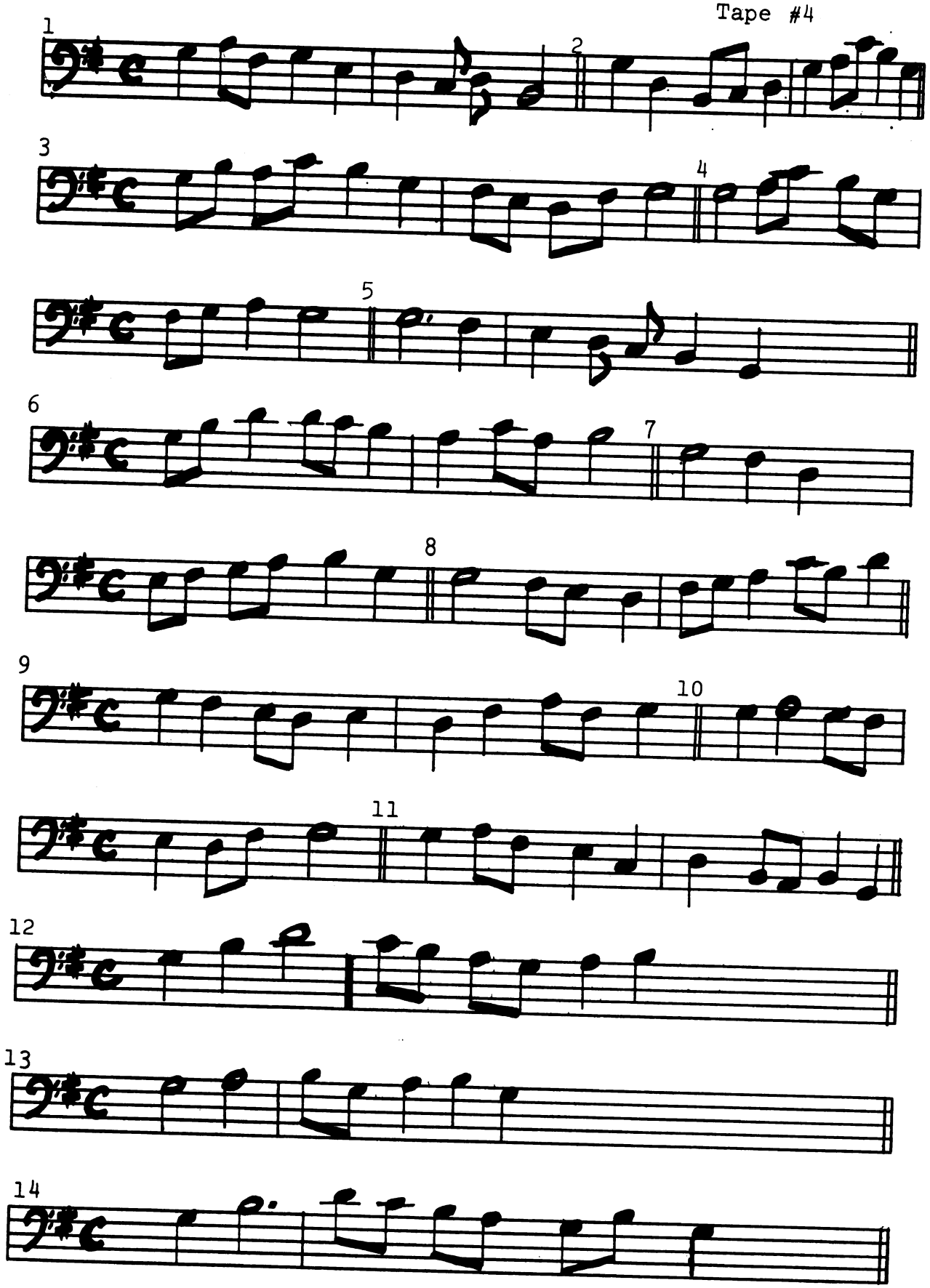
Series C
Tape #2



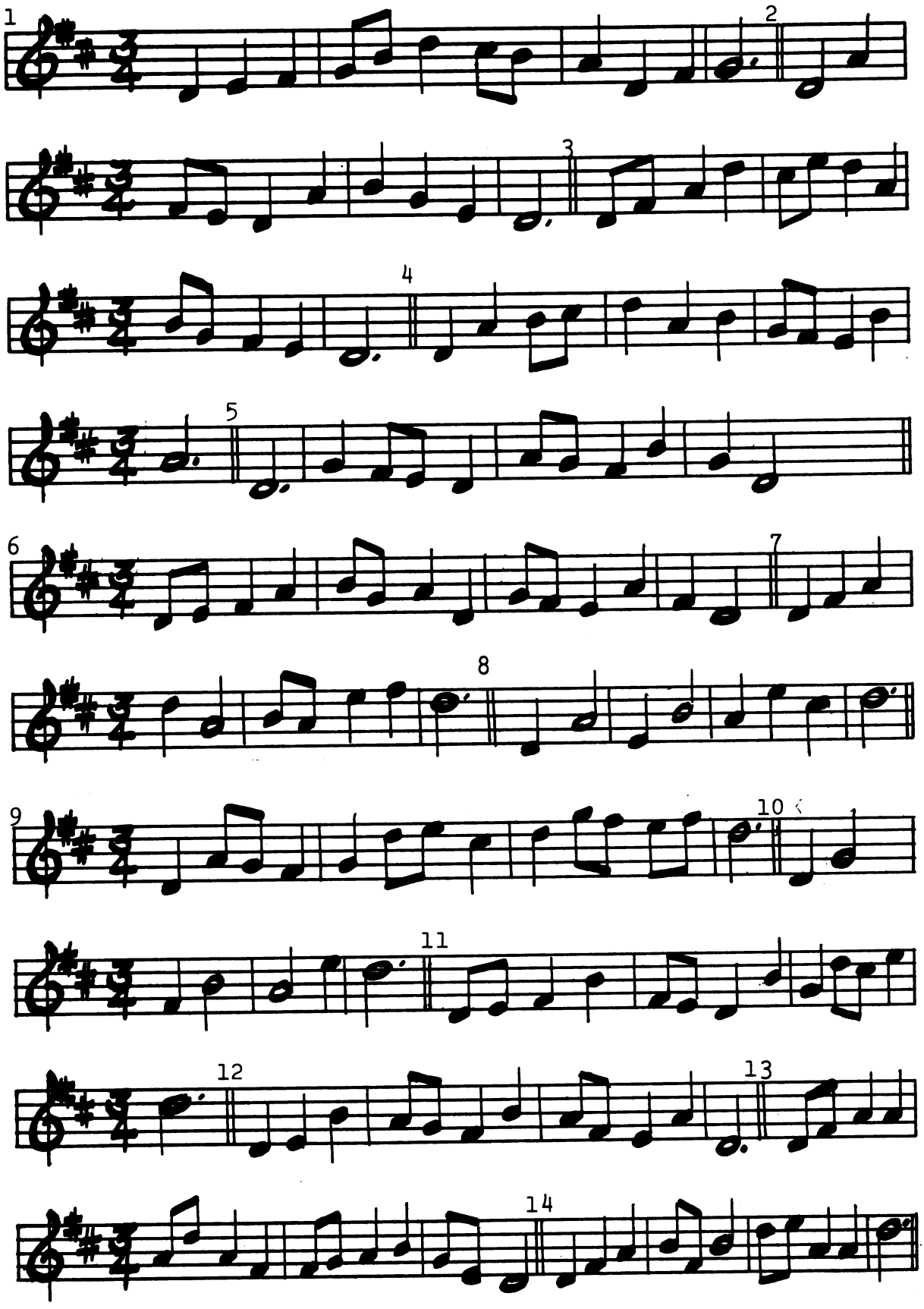
Series C
Tape #3



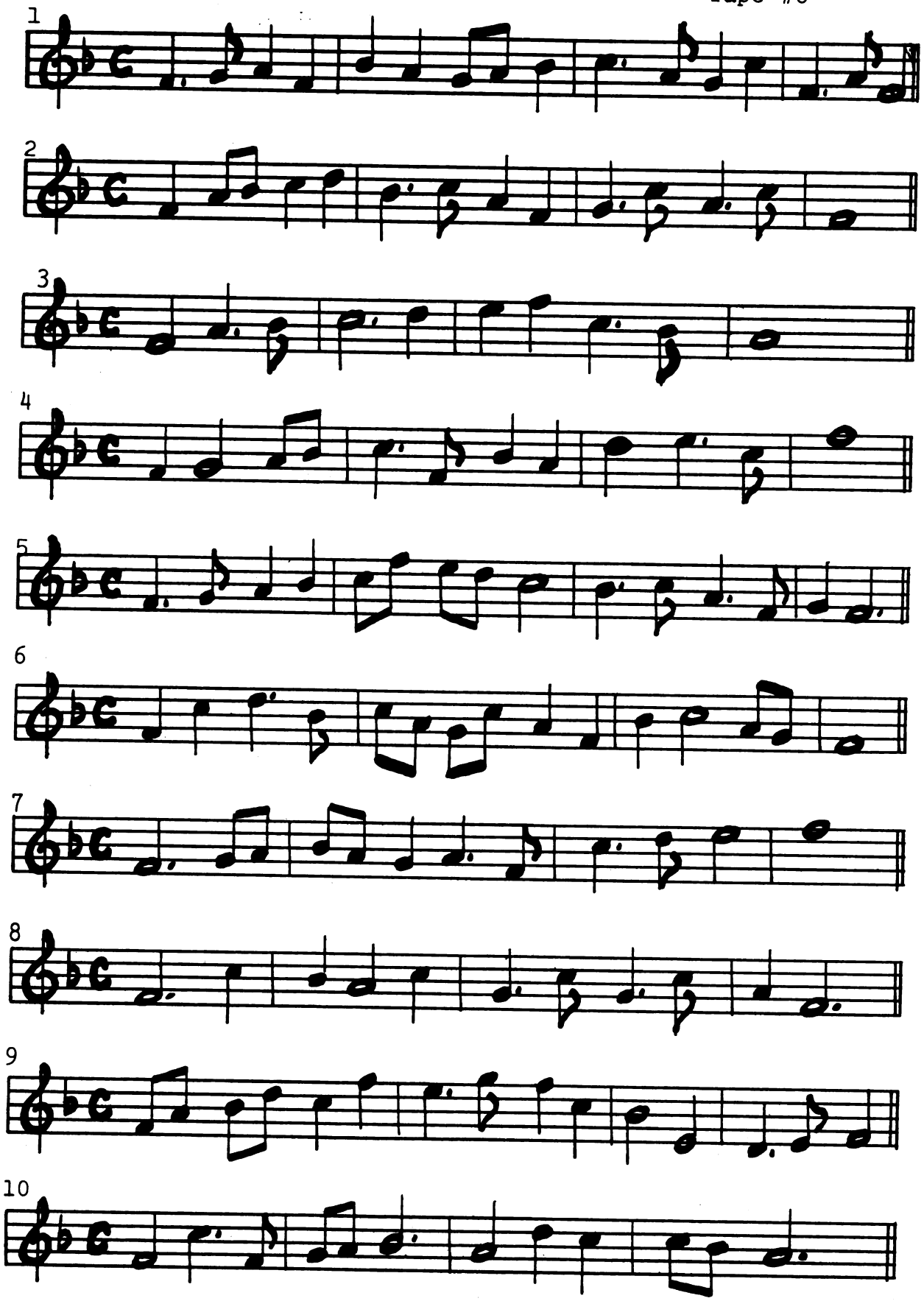
Series C
Tape #4



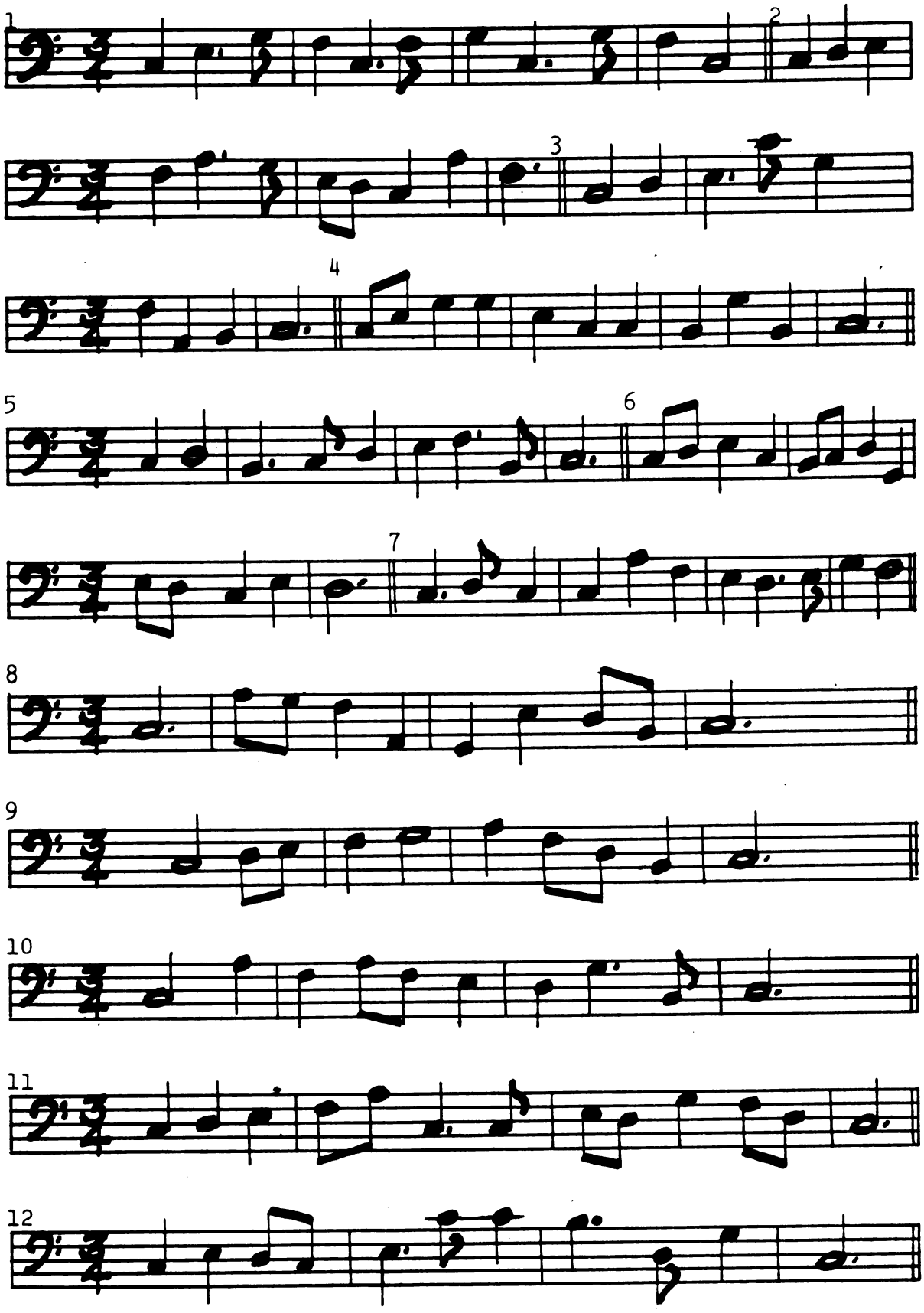
Series C
Tape #5



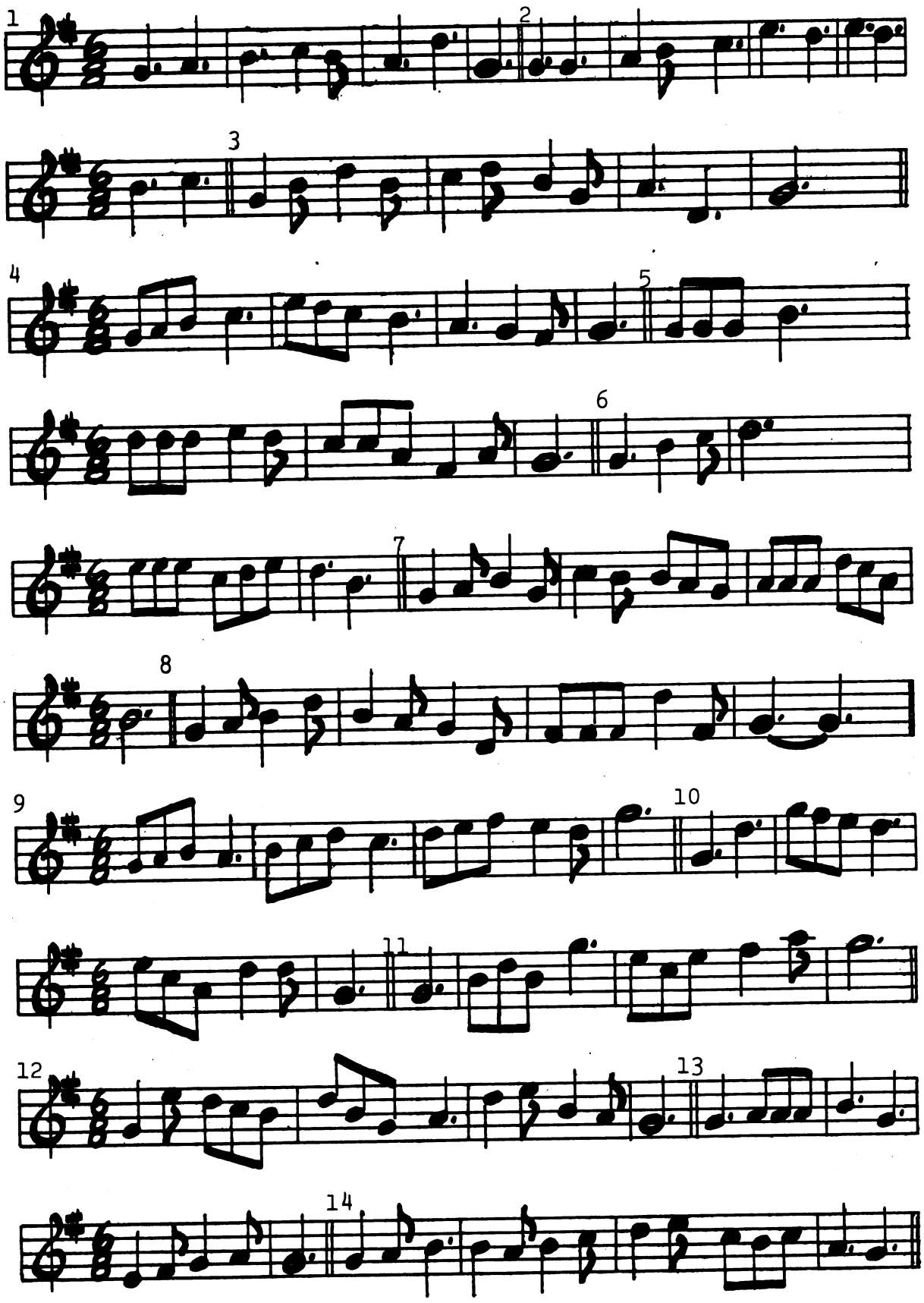
Series C
Tape #6



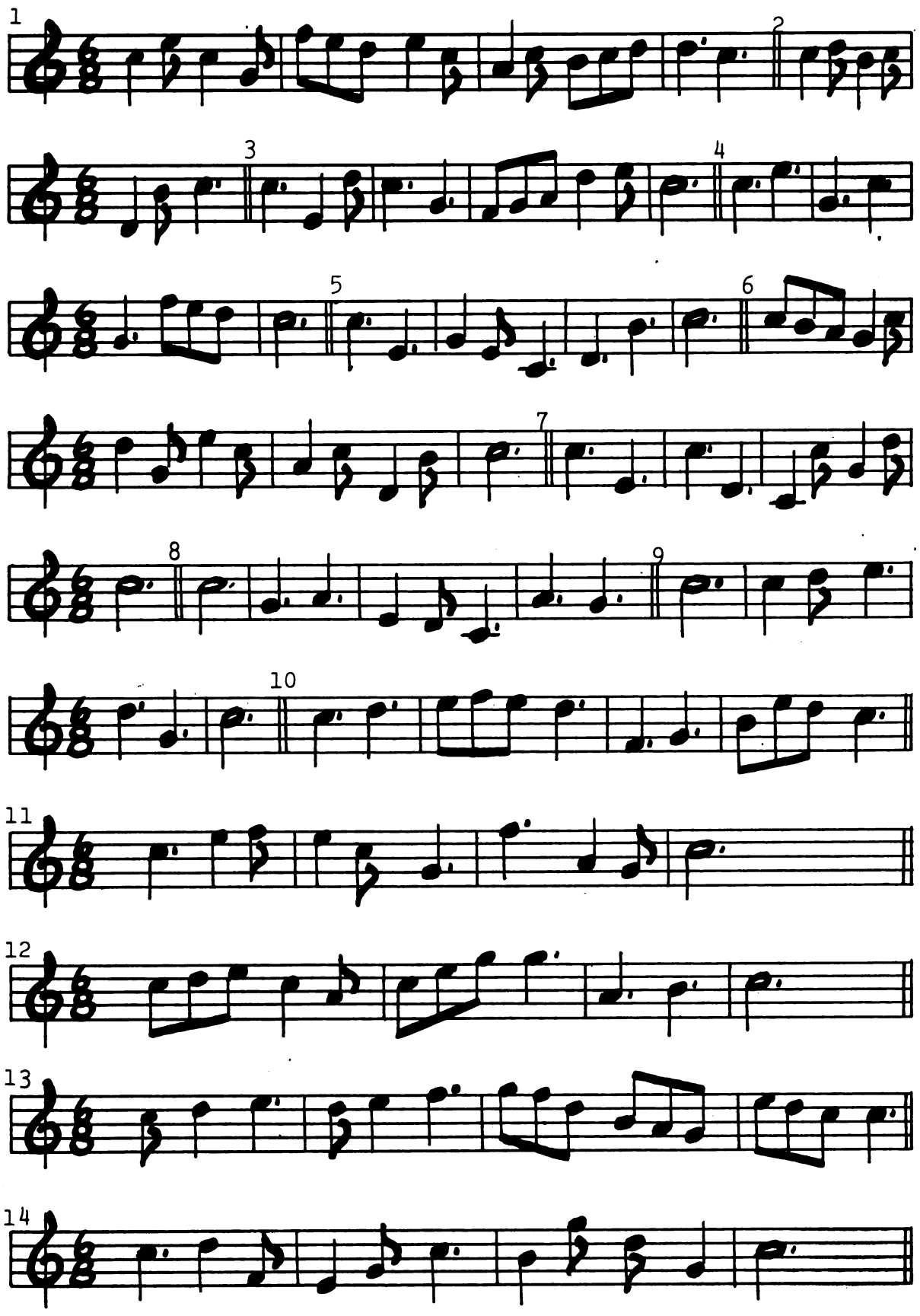
Series C
Tape #7



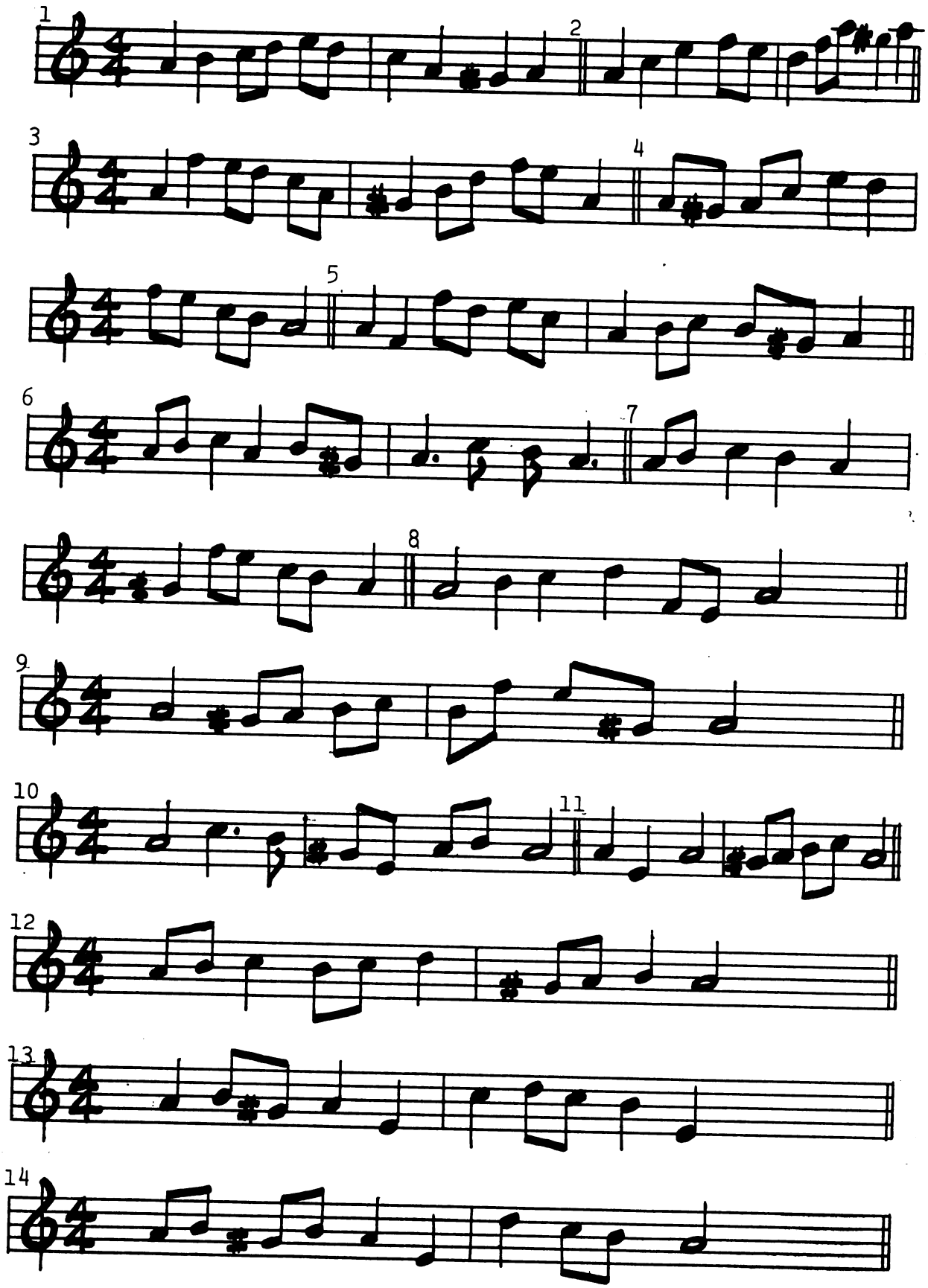
Series C
Tape #8



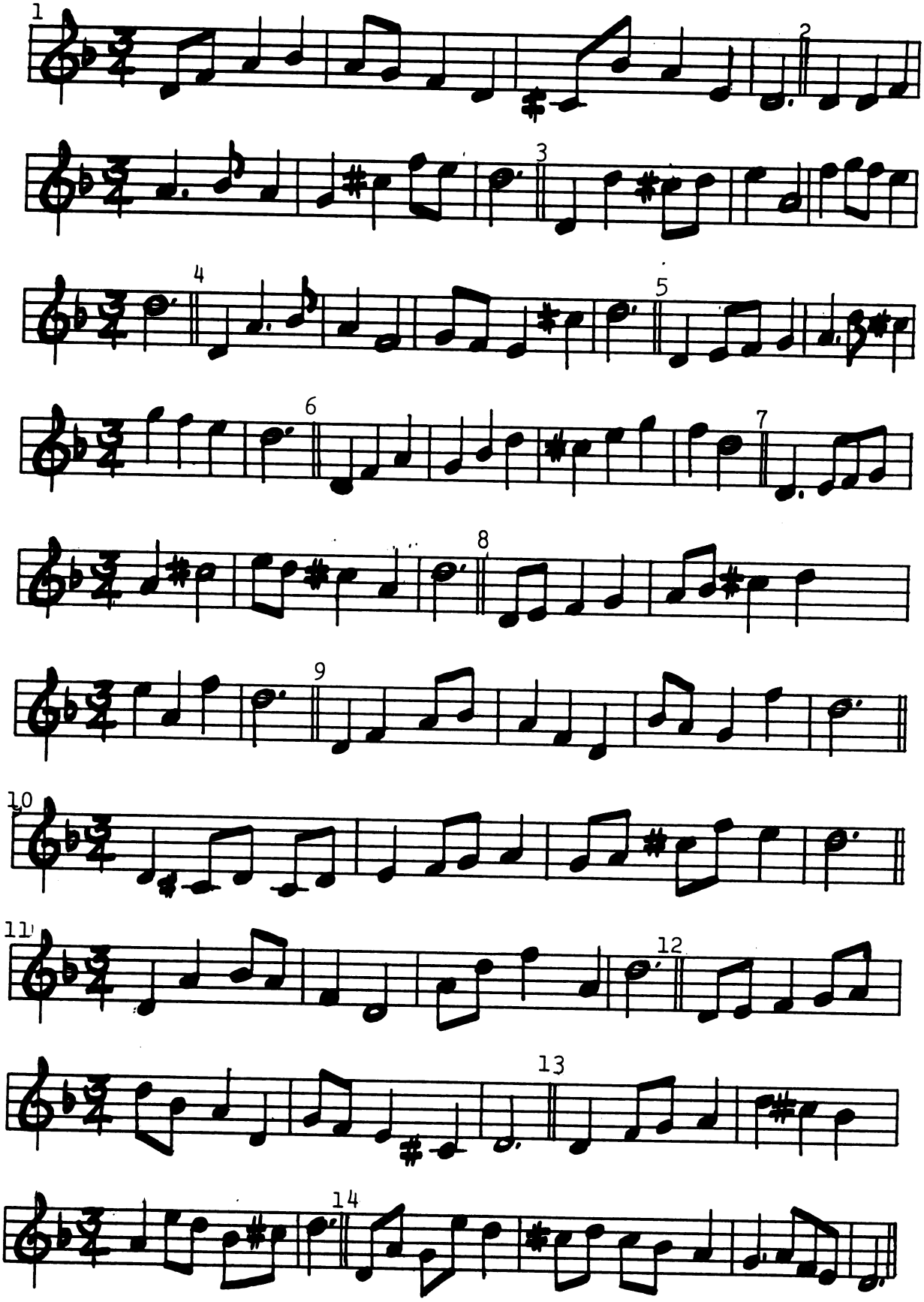
Series C
Tape #9



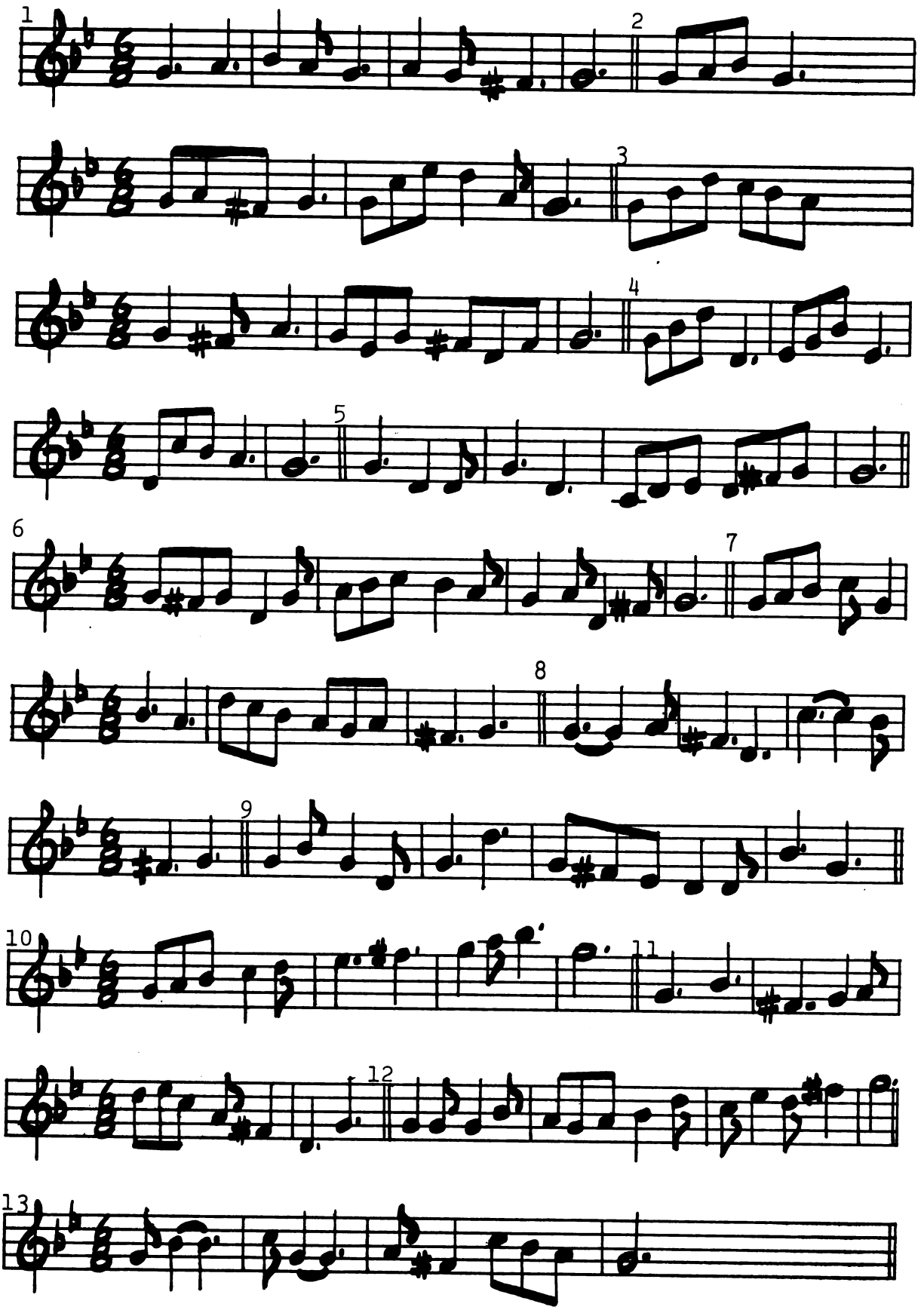
Series C
Tape #10



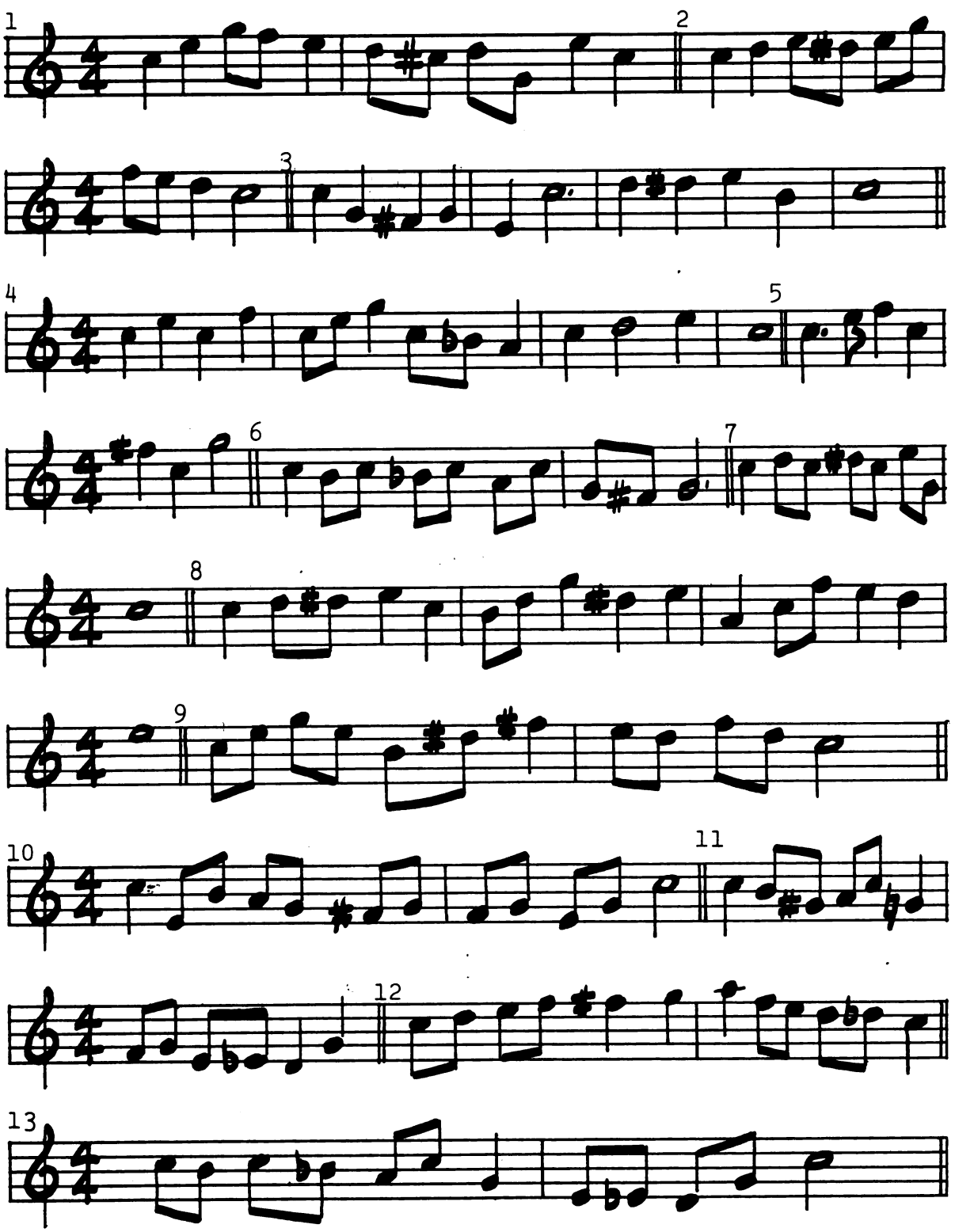
Series C
Tape #11



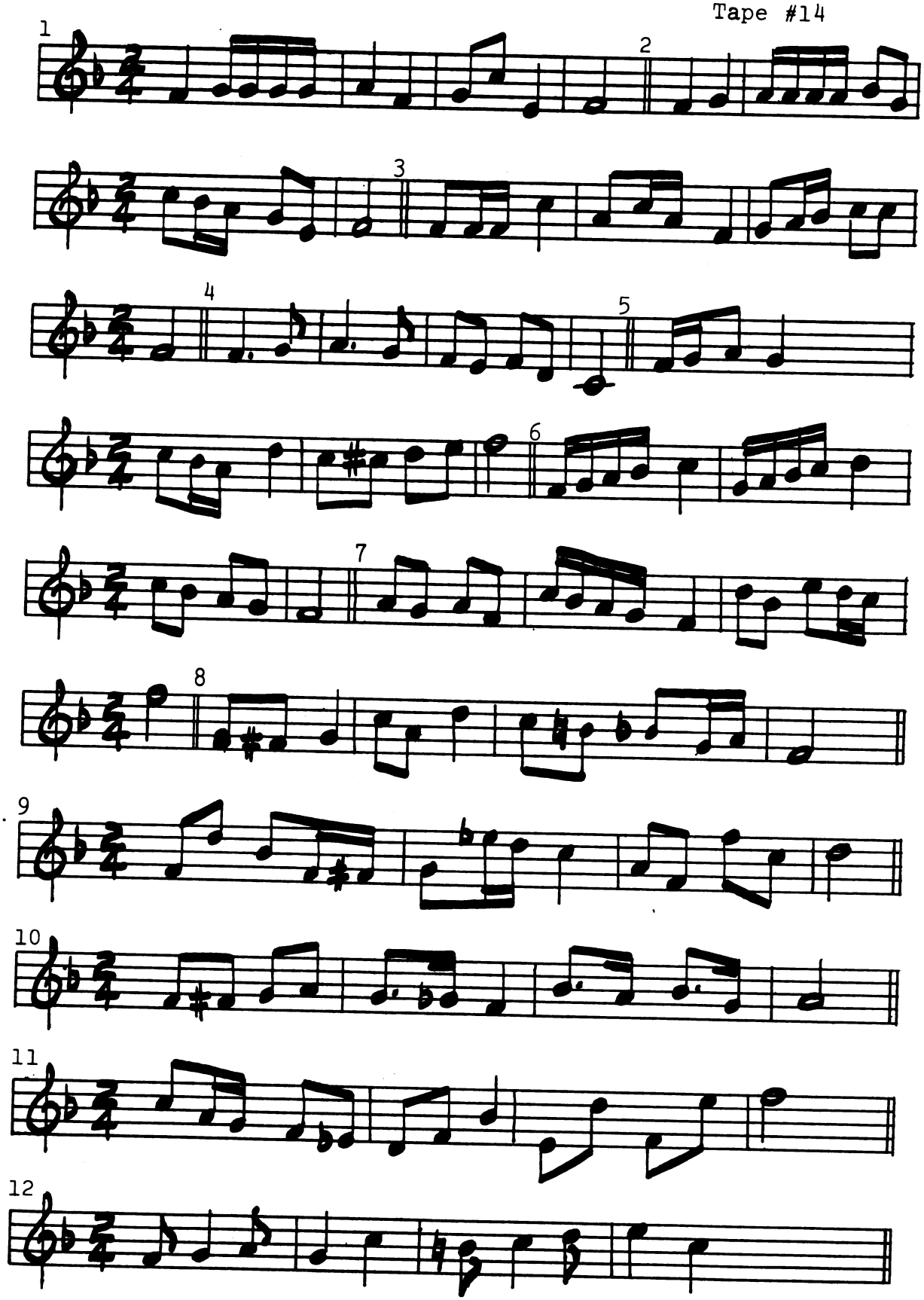
Series C
Tape #12



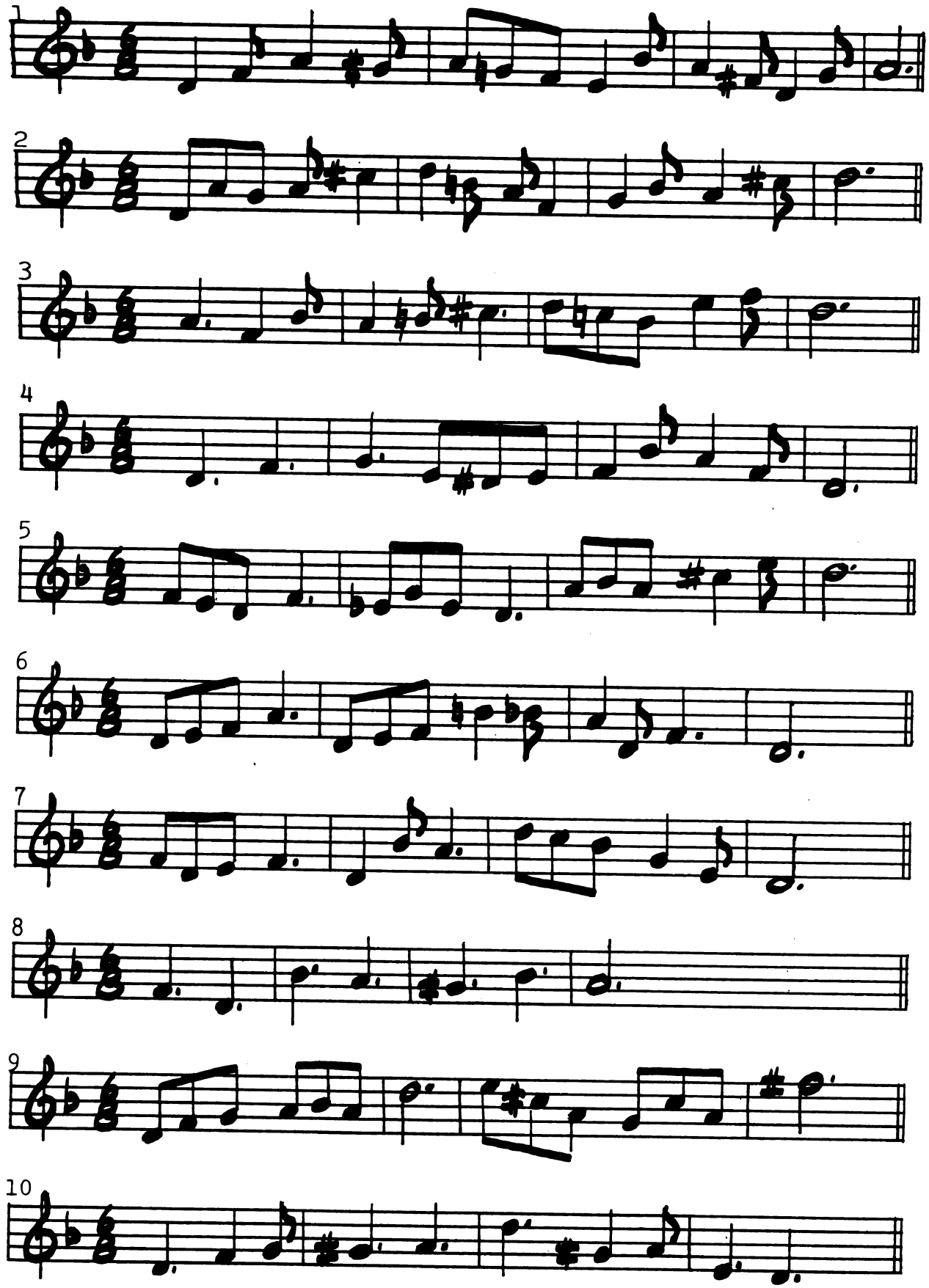
Series C
Tape #13



Series C
Tape #14



Series C
Tape #15



APPENDIX B

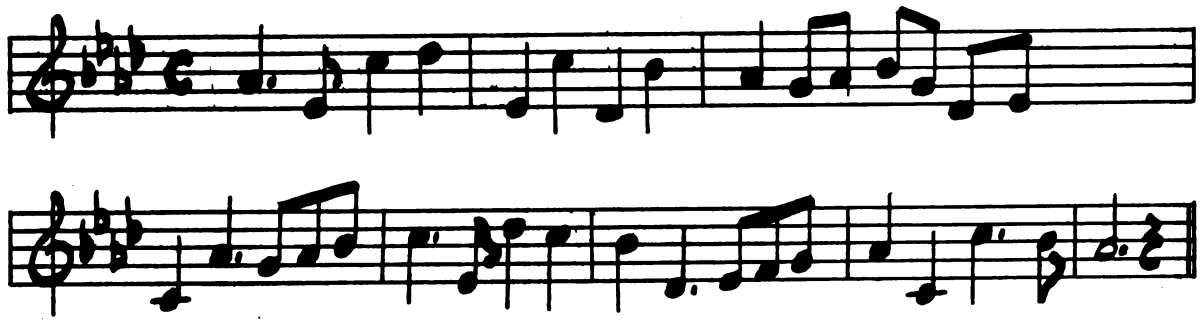
DICTATION AND SIGHT-SINGING TESTS

Sight-Singing Pre-Test

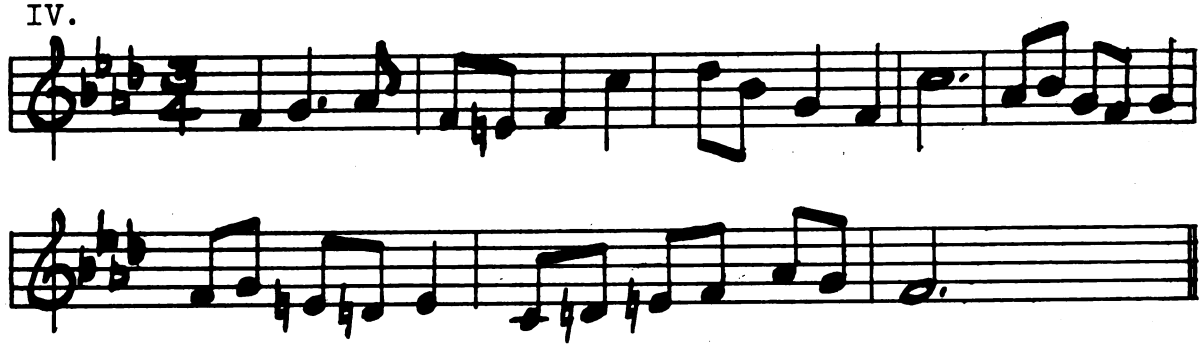
I. 1 2 3 5 6 4 7 3 1

II. 1 6 4 2 8 3 5

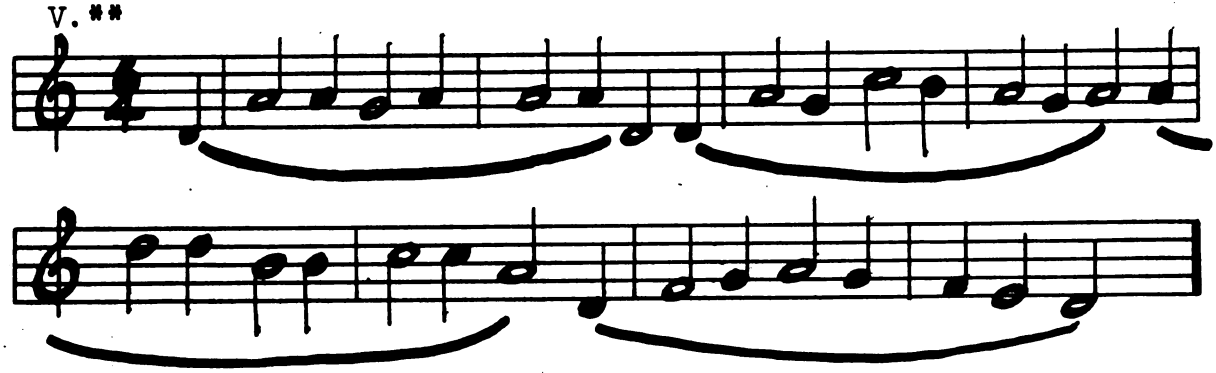
III.*



IV.



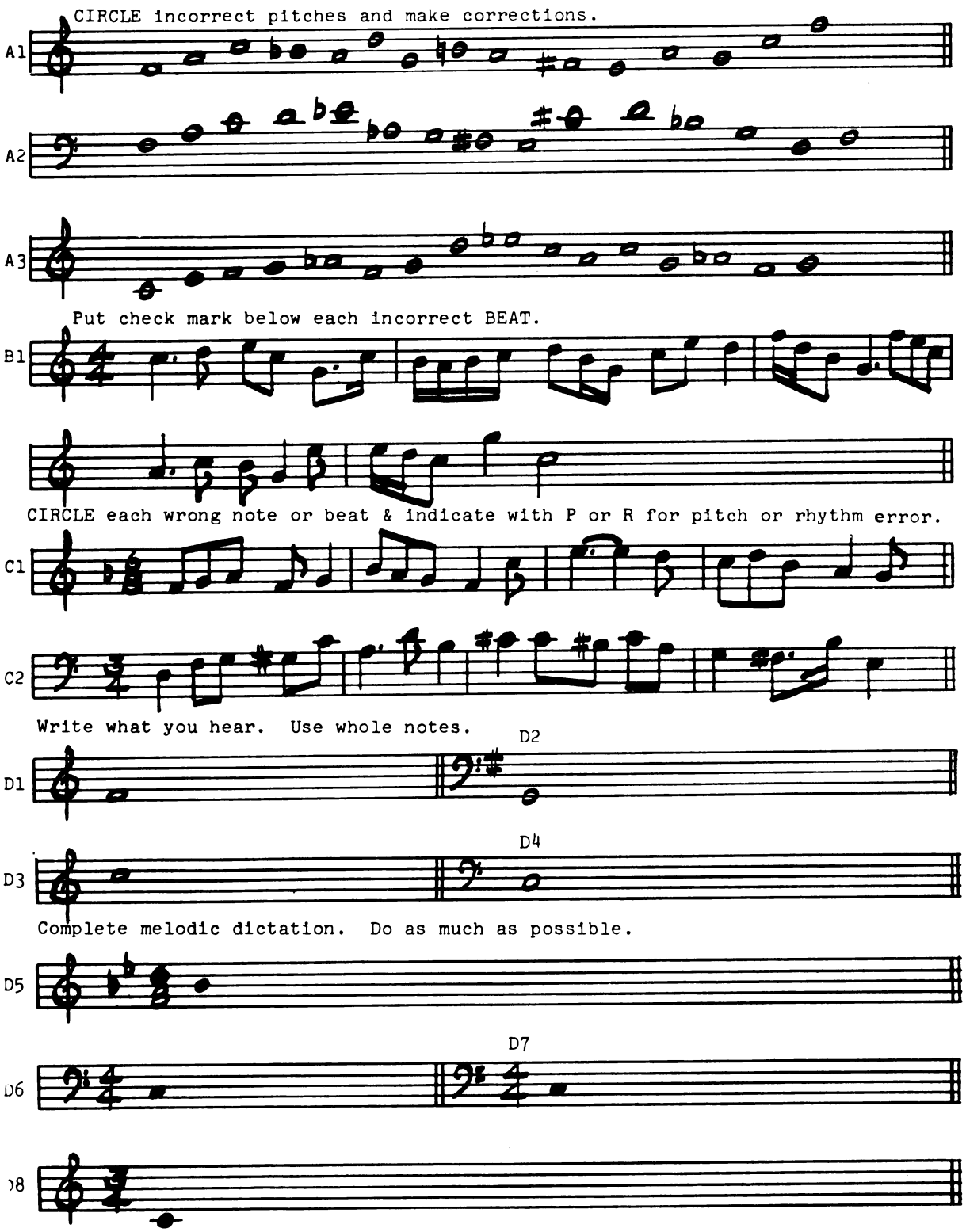
V.**



*Arnold Fish and Norman Lloyd, Fundamentals of Sight-Singing and Ear-Training (New York: Dodd, Mead and Company, 1964), p. 138, #4.

**Robert W. Ottman, Music for Sight-Singing (Englewood Cliffs, New Jersey: Prentice-Hall, 1956), p. 188, #20.

DICTATION TEST--STUDENT BLANK



DICTATION TEST