4psw26: Speech perception assessment and training system (SPATS-ESL) for speakers of other languages learning English

James D. Miller, Roy Sillings, Charles S. Watson, & Diane Kewley-Port
Communication Disorders Technology, Inc.
510 N. Morton Street, Suite 215
Bloomington, IN 47404

Abstract

The SPATS software system (1,2, 3, & 4), originally developed for the hearing impaired, has been modified for use with ESL learners with TOEFL (pbt) scores near and well above 500. SPATS-ESL includes the identification of syllable constituents: onsets, nuclei, and codas as well as sentence recognition. The syllable constituent tasks include the progressive introduction of increasing numbers of constituents until the learner becomes adept at the identification of 45 onsets, 28 nuclei, and 36 codas spoken by eight talkers in a variety of phonetic contexts. The sentence task emphasizes increasing speed and decreasing errors in the recognition of short, meaningful, sentences spoken naturally by a variety of talkers. The sentences are presented in a background of multi-talker babble at five signal-to noise ratios: +5, 0, -5, -10 & -15 dB. The syllable constituent and sentence tasks are interleaved throughout training. In constituent training, SPATS uses a proprietary training algorithm, Adaptive Item Selection (AIS), that focuses training on items of that are of intermediate difficulty for each individual learner. Proctored tests allow certification of the level of a learner's English speech perception skills relative to those of native speakers.

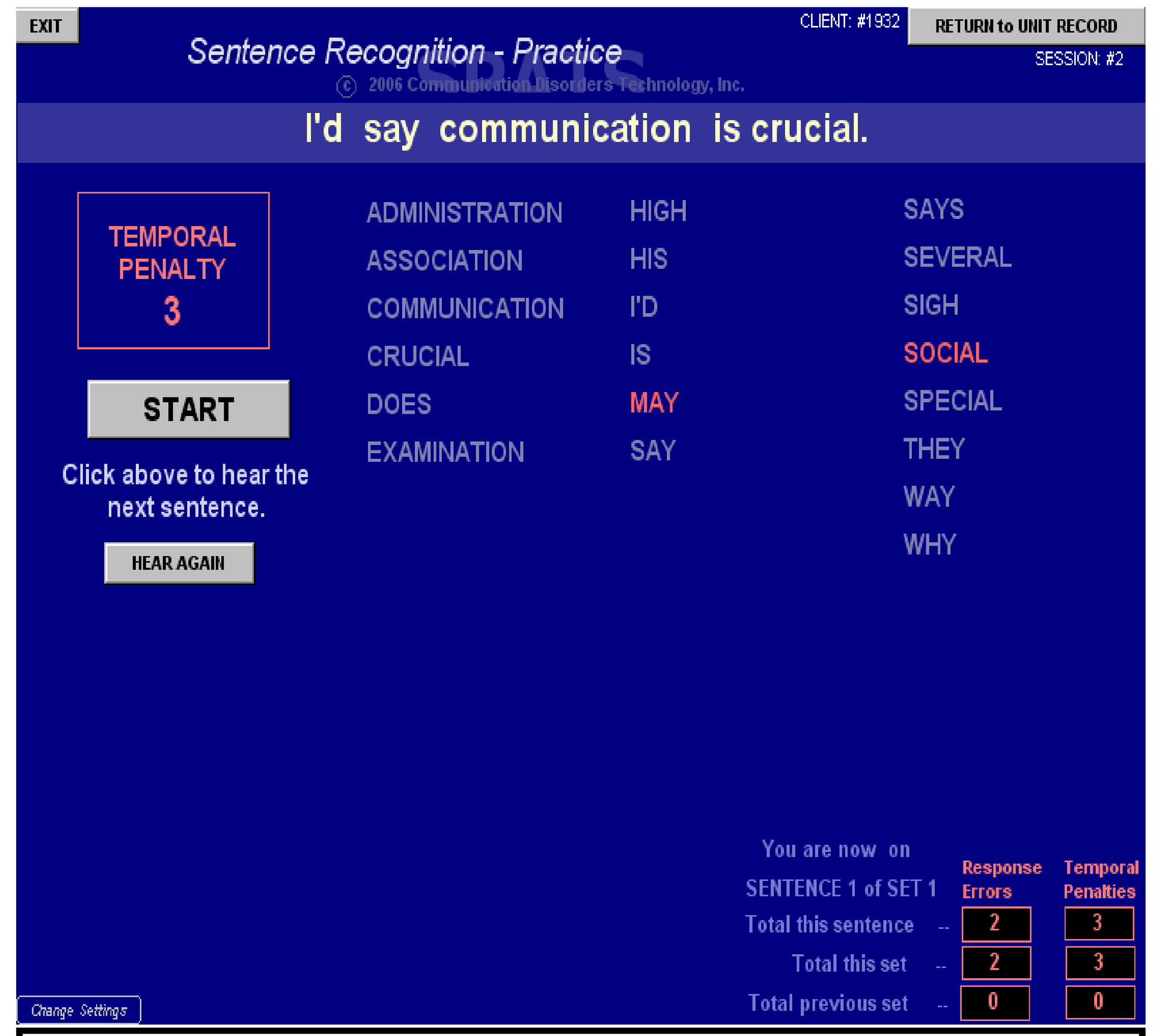
Modules: SPATS-ESL has four modules:
I. The Sentence Module,
II. The Syllable-Constituent Module,
III. The Report Module,
and
IV. The Proctored Test Module.

I. The Sentence Module

The sentence module provides practice in top-down and combined top-down and bottom up speech perception skills. One thousand sentences have been recorded by ten different talkers. Each is spoken naturally and the rate of speech, intonation patterns, and stress patterns vary among the talkers and sentences. Therefore, the range of phonetic accommodations that occur in everyday speech are found in this corpus.

Scoring of the sentence task is objective and entirely computer based.

The basic task: A spoken sentence of three to seven words is presented. A screen then appears that shows "slots" for each word at the top and an alphabetical list of words that contains the spoken words plus three phonetically similar foils for each. The user is instructed to click on the words that they thought they heard. Correctly selected words are "dimmed" and appear in the appropriate slot in the header. Errors turn red and cause the sentence to be replayed. Whenever the listener goes 5 sec. without a response a "temporal penalty" is assessed. Beginning and final screens for a five word sentence are shown next.



In each group of 15 sentences, three are at each of 5 SNRs: +5, 0, -5, -10, & -15 dB. The learner is shown his overall effective percent correct after the completion of each group of 15 sentences. The effective percent correct is the total number of words divided by the number words plus the number of errors plus the number of temporal penalties times 100. One learner's progress is shown in the next panel. Native speakers score 90 and above.

Sentences	Words	Errors	Temporal Penalties	Effective Percent Correct	Performance Category
1-15	84	20	20	68	Needs Work
16-30	85	16	12	75	Fair
31-45	82	10	8	82	Fair
46-60	87	5	8	87	Very Good
61-75	85	9	9	83	Good

II. The Syllable-Constituent Module

- Constituent Types: Onsets, Nuclei, and Codas
- Constituents are selected in preference to phonemes so that allophonic variations are well represented in testing and training.
- Constituents are ordered in importance based on their average ranks in lexical and textual frequency of occurrence.
- SPATS features the **post-response rehearing option** (ask for explanation).
- Benchmark Testing and Training progresses from Cumulative Level 1 to Cumulative Level 4 (see columns to the left).
- Based on a listener's performance every item in a set has a current **Item Mastery Score (IMS)** of 100-very easy, 75-easy, 50-moderate, 25-difficult, or 0-very difficult.
- When an item is correctly identified, its IMS increases 25 points.
- When an item is missed or incorrectly used as a response, its IMS decreases 25 points. (No score can go above 100 or below 0.)
- In Benchmark Tests items are presented equally often, and all items begin with an IMS of 50.
- In Training Runs an item's IMS is brought forward from the previous Benchmark Test or Training Run.
- For any Constituent Type and Level, there are 4 Training Runs between Benchmark Runs.
- In Training Runs items are presented based on their IMS's by the Adaptive Item Selection (AIS) algorithm.
- Adaptive Item Selection (AIS) is a Markov process that focuses training on items of Moderate Difficulty as indicated by the table.

Difficulty Category	Probability of Selection
100-very easy	0.100
75-easy	0.200
50-moderate	0.400
25-diffiuclt	0.200
0-very difficult	0.100
Sum of Prob.	1.000

List of Onsets

Col. 1		Col. 2		Col. 3		Col. 4	
1	Vwl	12	p-	24	thv-	35	dr-
2	s-	13	l-	25	br-	36	kw-
3	r-	14	n-	26	gr-	37	fl-
4	k-	15	pr-	27	kr-	38	sl-
5	b-	16	g-	28	sp-	39	kl-
6	m-	17	v-	29	pl-	40	sw-
7	h-	18	st-	30	th-	41	sm-
8	d-	19	sh-	31	y-	42	gl-
9	t-	20	fr-	32	sk-	43	sn-
10	w-	21	j-	33	bl-	44	skr-
11	f-	22	tr-	34	str-	45	z-
		23	ch-				

The number next to each onset is its importance based on its average rank in lexical and textual frequency of occurrence

Each onset is spoken by 8 talkers and combined with 4 vowels: "ee," "ah," "oo," & "er."

Cumulative Level 1 = Col. 1
Cumulative Level 2 = Col. 1 + 2
Cumulative Level 3 = Col. 1 + 2 + 3
Cumulative Level 4 = Col. 1 + 2 + 3 + 4

List of Nuclei

Col. 1		Col. 2		Col. 3		Col. 4	
2	heed	8	hayed	1	hid	5	head
3	heard	12	hulled	6	hud	10	hawed
4	had	13	hoed	14	hoard	15	how'd
7	hod	16	hewed	18	haired	23	Hal'd
9	hide	17	hard	19	hood	24	hold
11	whoed	20	hilled	26	hired	25	hauled
22	held	21	hoid	28	heeled	27	hailed

The number next to each nucleus is its importance based on its average rank in lexical and textual frequency of occurrence

Cumulative Level 1 = Col. 1
Cumulative Level 2 = Col. 1 + 2
Cumulative Level 3 = Col. 1 + 2 + 3
Cumulative Level 4 = Col. 1 + 2 + 3 + 4

Ask for explanation of grouping of nuclei.

List of Codas

Col. 1		Col. 2		Col. 3		Col. 4	
1	Vwl	10	-m	19	-bz	28	-ch
2	-z	11	-nt	20	-ks	29	-sts
3	-n	12	-v	21	-p	30	-pt
4	-d	13	-k	22	-dz	31	-ngz
5	-t	14	-nz	23	-mz	32	-vz
6	-l	15	-ts	24	-f	33	-ps
7	-ng	16	-r	25	-kt	34	-sh
8	-s	17	-st	26	-th	35	-g
9	-nd	18	-ns	27	-j	36	-b

The number next to each coda is its importance based on its average rank in lexical and textual frequency of occurrence

Each coda is spoken by 8 talkers and combined with 5 stems: "ee," "ah," "oo," "er," & "el."

Cumulative Level 1 = Col. 1
Cumulative Level 2 = Col. 1 + 2
Cumulative Level 3 = Col. 1 + 2 + 3
Cumulative Level 4 = Col. 1 + 2 + 3 + 4

Opinion: It is the requirement of attending to all 1998 contrasts that induces the ESL learner to learn the categories and dimensions of the sound system of English. This is consistent with Kingston's (5) view that the ESL learner must learn to attend to the dimensions of the English sound system. This lays a foundation for accurate, rapid perception of spoken English and provides a necessary foundation for the acquisition of correct pronunciation and accent reduction.

Items and Trials per Run for Levels 1 - 4

Cumulative Level	Number Onsets	Onset Contrasts	Trials per Item	Total Trials
1	11	55	4	44
2	23	253	4	92
3	34	561	4	136
4	45	990	4	180
4 Maint.	45	990	2	90

Items and Trials per Run for Levels 1 - 4

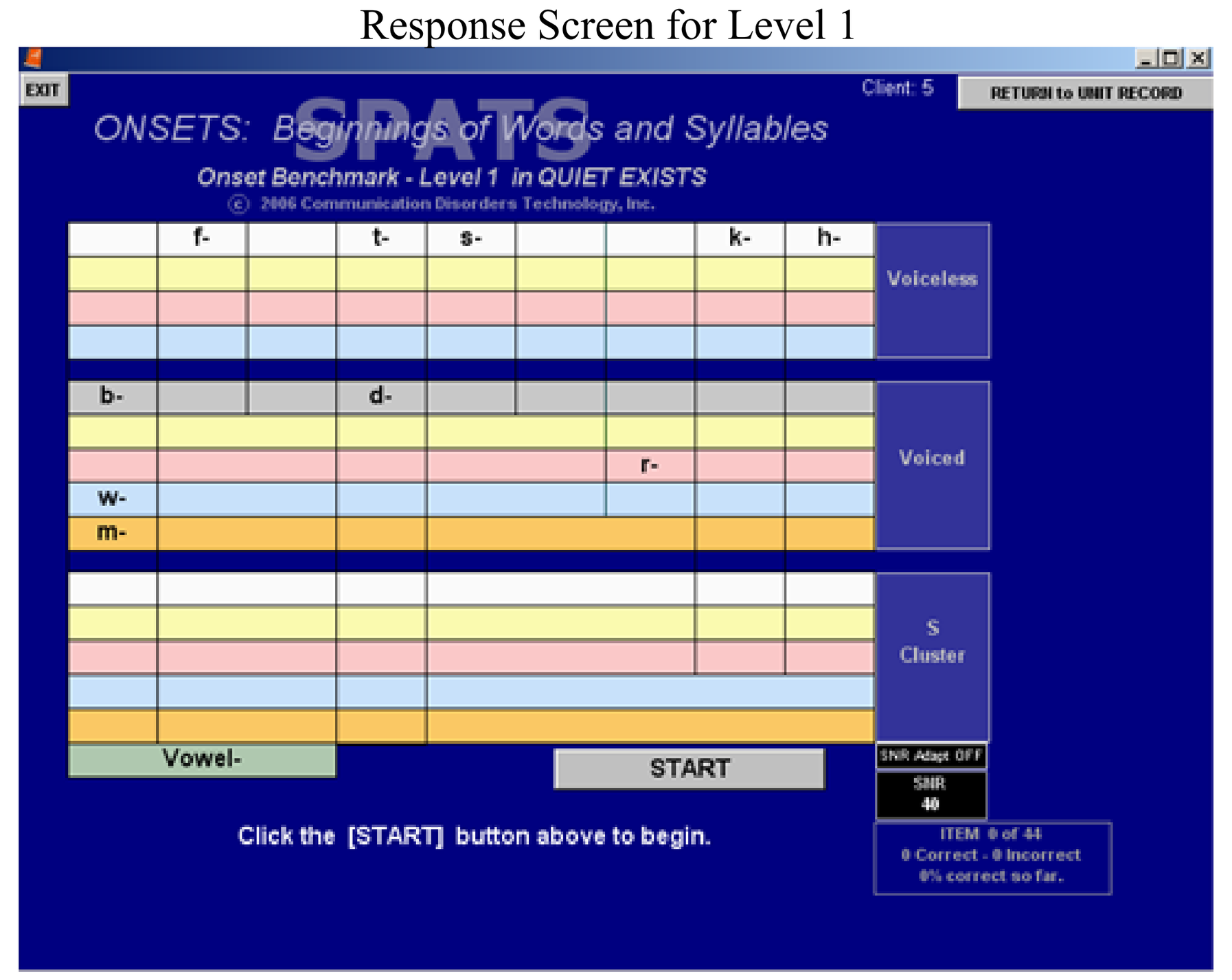
Cumulative Level	Number Nuclei	Nucleus Contrasts	Trials per Item	Total Trials
1	7	21	7	49
2	14	91	7	98
3	21	210	7	147
4	28	378	7	196
4 Maint.	28	378	2	56

Items and Trials per Run for Levels 1 - 4

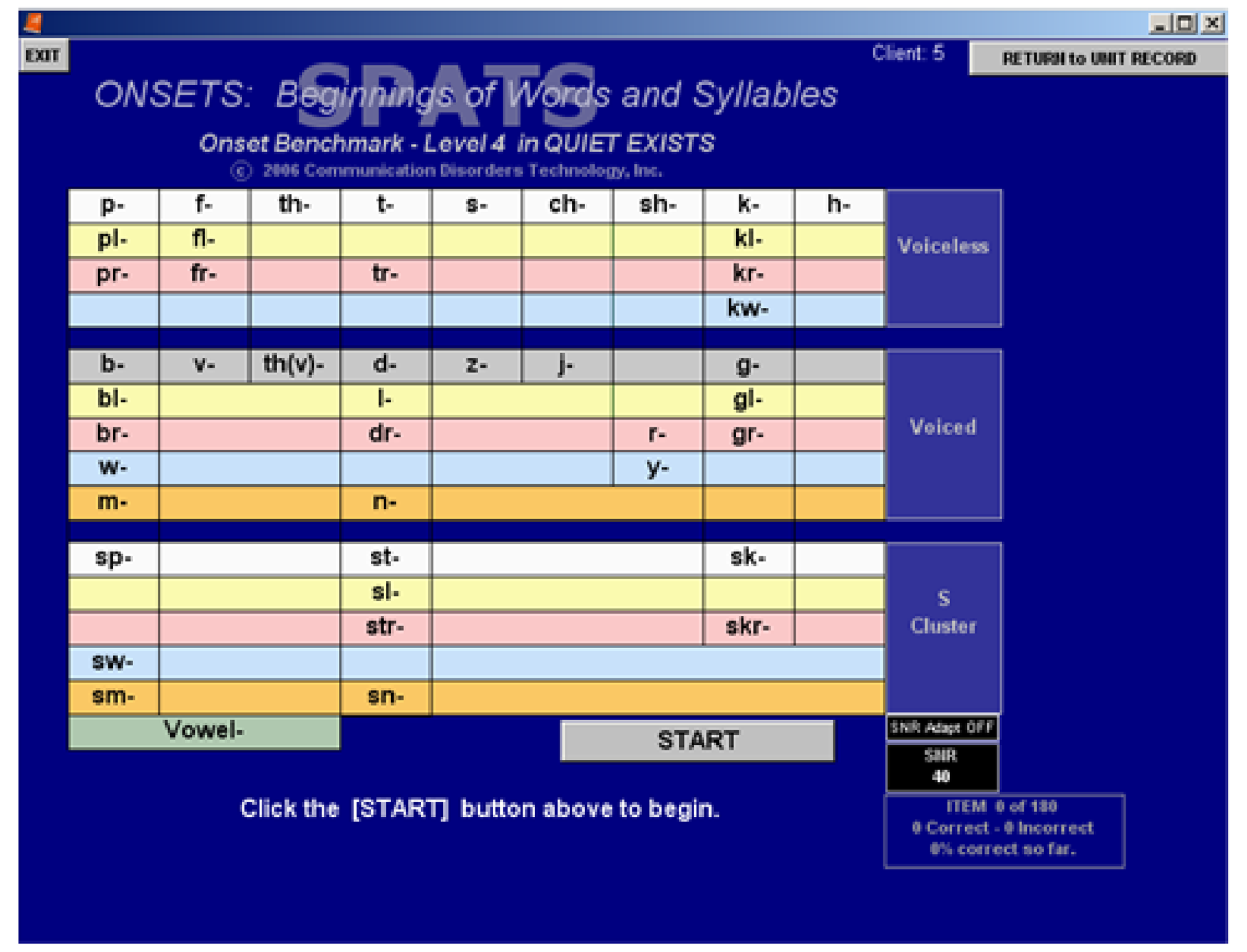
Cumulative Level	Number Codas	Coda Contrasts	Trials per Item	Total Trials
1	9	36	5	45
2	18	153	5	90
3	27	351	5	135
4	36	630	5	180
4 Maint.	36	630	2	72

Cumulative Level	Total Constituents	Total Contrasts	Total Trials
1	27	112	138
2	55	497	280
3	82	1122	418
4	109	1998	556
4 Maint.	109	1998	218

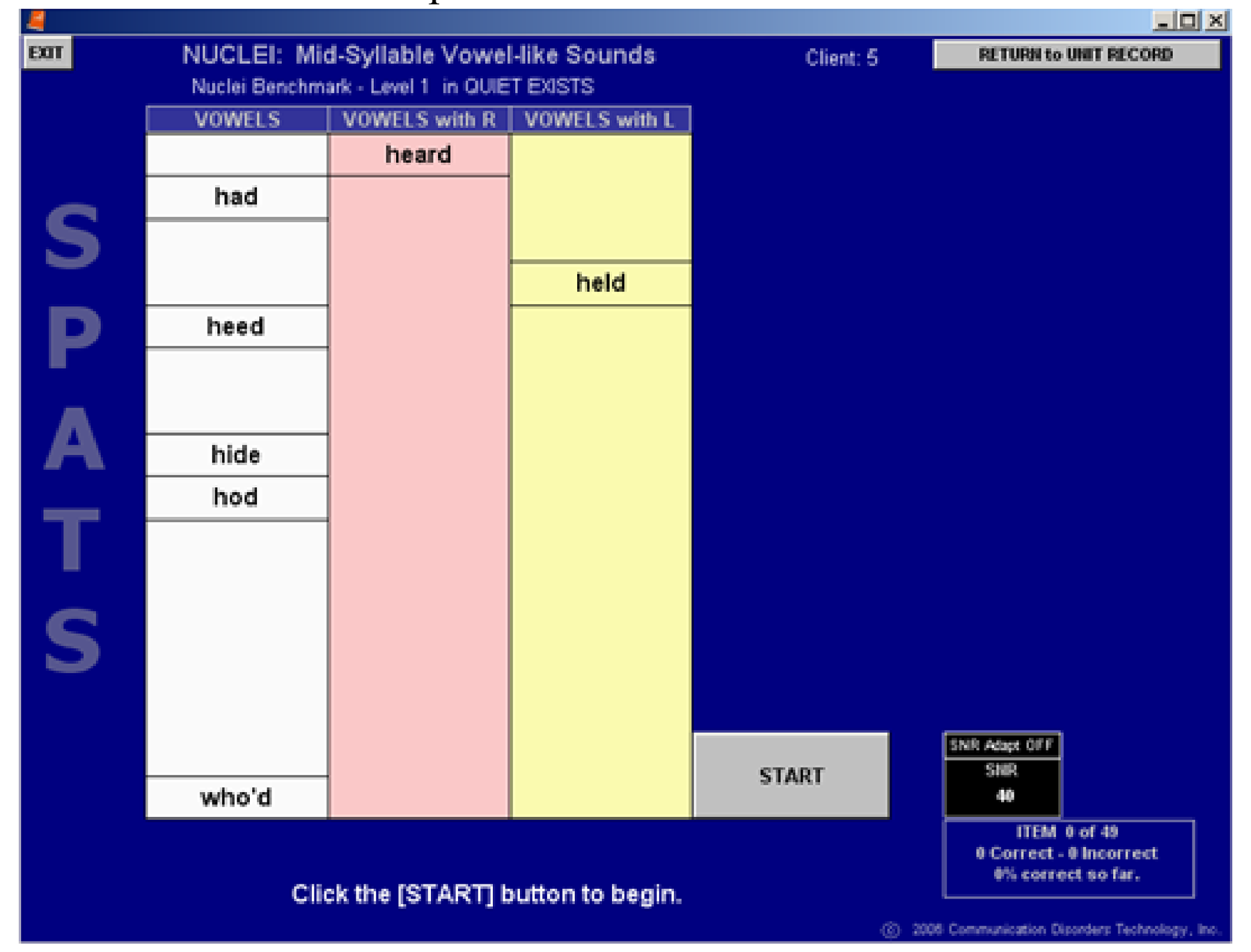
Response Screen for Level 1



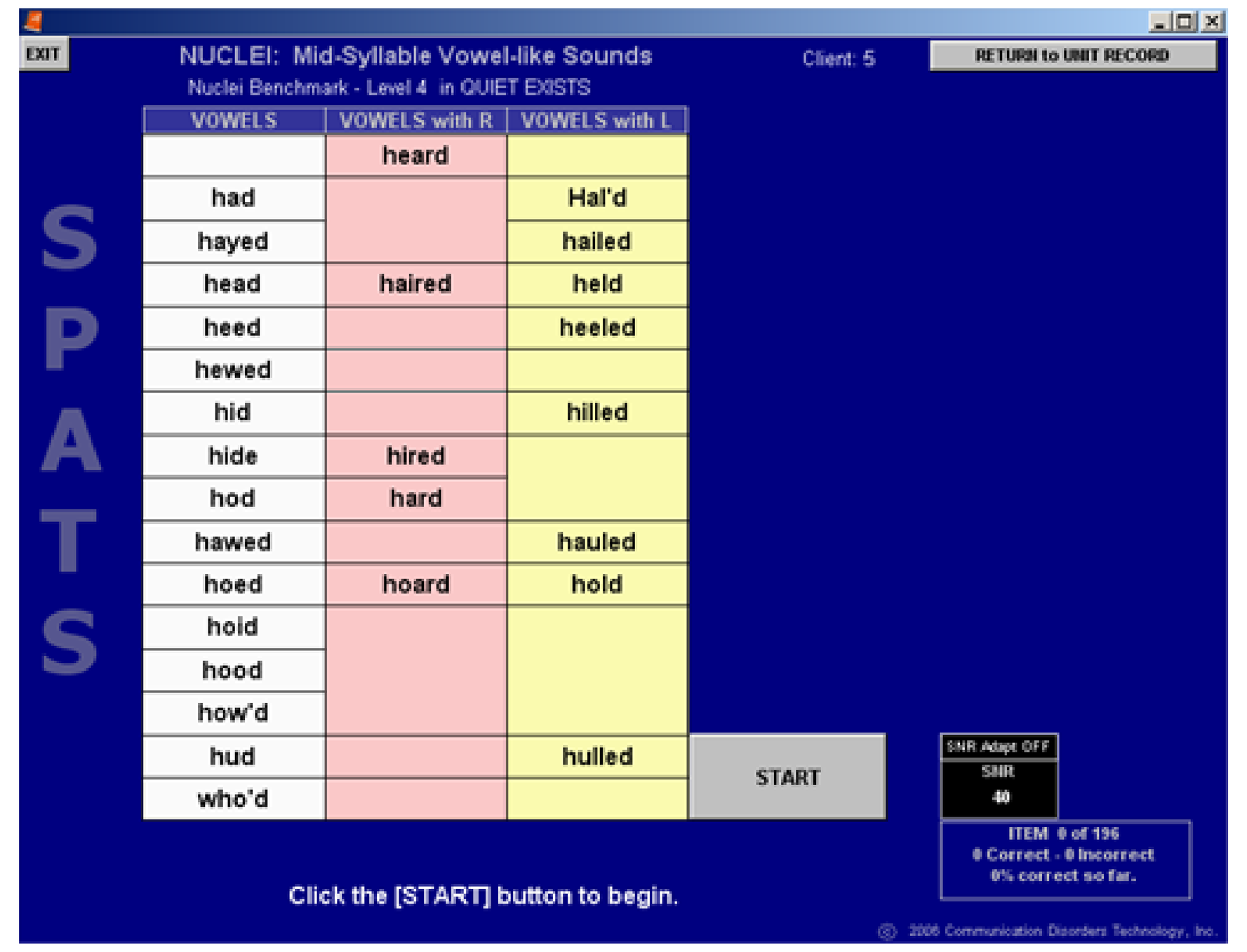
Response Screen for Level 4



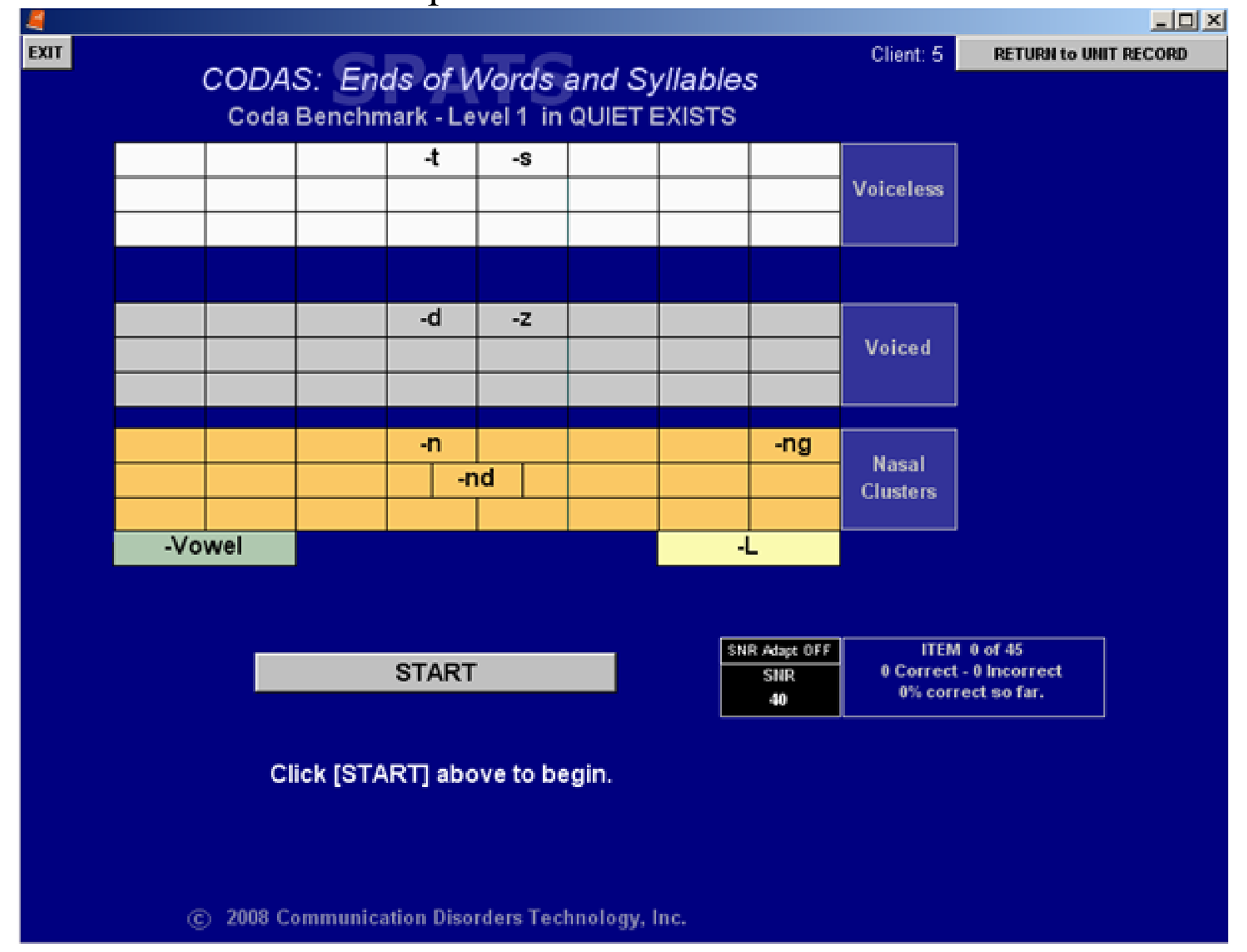
Response Screen for Level 1



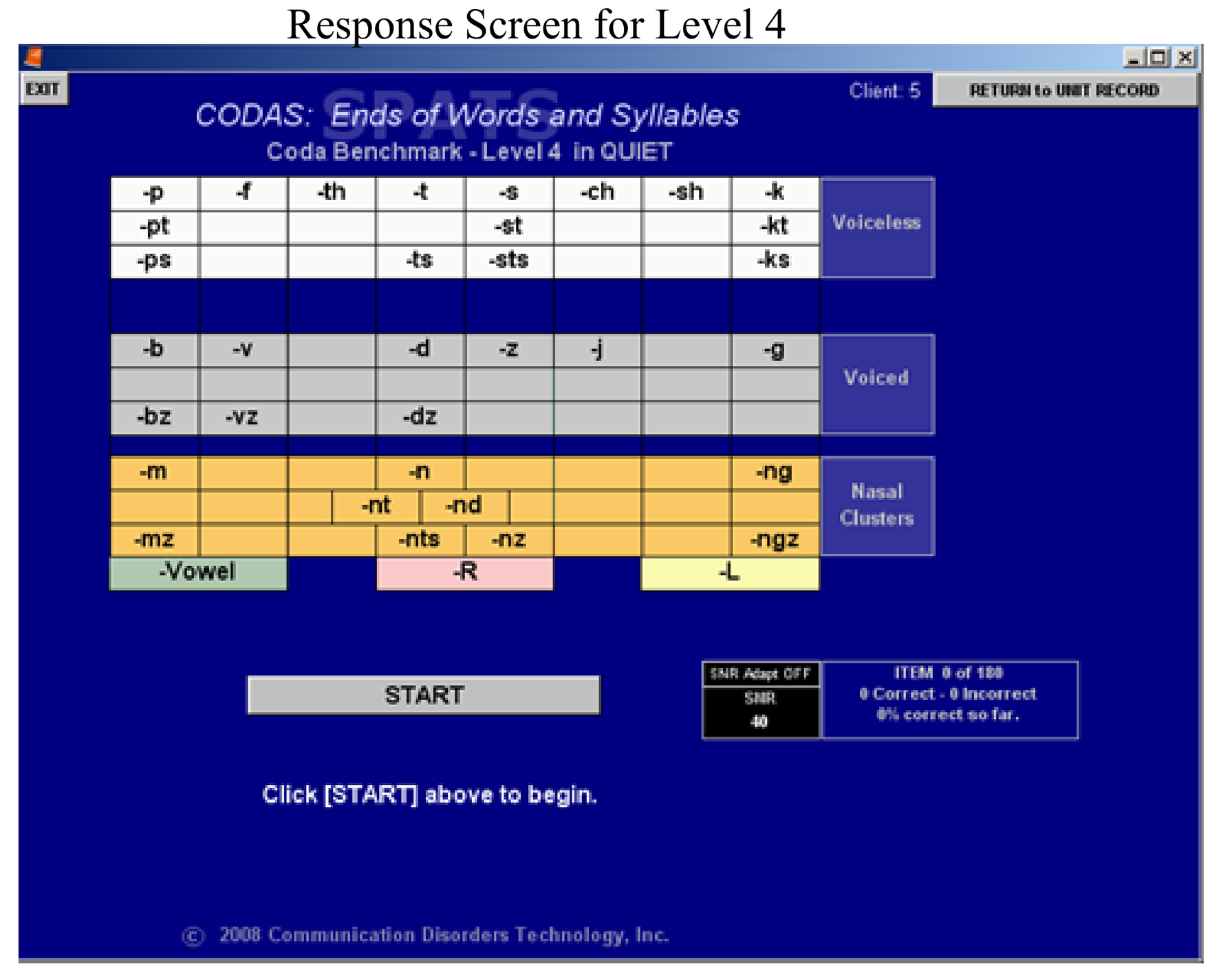
Response Screen for Level 4



Response Screen for Level 1



Response Screen for Level 4



System Requirements

SPATS-ESL runs on PCs with XP OS. Headsets comparable to the Sennheiser HD 212 are required.
If internet access is available, client records are kept on a secure web-server and the client may work at any SPATS-ESL equipped PC with XP OS. The program will automatically access a client's records and download them to the current machine and return them to the web server at the end of the session.

The Default Curriculum and How it Works

1) The software automatically gives detailed on-screen instructions and introductions to each new task.

2) All tasks are presented automatically by the software. The learner simply clicks a button labeled "CLICK HERE to CONTINUE," and follows on-screen instructions.

3) The default curriculum is "progressive" starting at Level 1 and progressing to higher levels as the learner reaches criteria. The criterion for advancement is that the learner must approach or exceed the lower limit of native speaker performance.

4) Criteria for advancement can be met during either Benchmark or Training Runs.

5) Each new Level begins with a Benchmark Run with 4 Training Runs per Benchmark Run, thereafter.

6) Clients are urged to work in sessions of 20-90 minutes per day and to schedule 3-5 sessions per week.

7) A User's Guide (in final editing) provides supplemental information and will be available by mid June at **www.comdistec.com.**

The Default Curriculum

-- PROGRESSIVE at 100% Native IMS
Sentences INTRO
Sentences PRETEST no audio
Sentences PRETEST with audio
-- Begin Rotation #1 (on 0 of 100)
Onsets L1 INTRO ii
Onsets L1 QUIET i
Sentences PRACTICE 2 sets of 3
Nuclei L1 INTRO ii
Nuclei L1 QUIET i
Sentences PRACTICE 2 sets of 3
Codas L1 INTRO ii
Codas L1 QUIET i
Sentences PRACTICE 1 set of 3
-- End Rotation #1 -----
-- CURRICULUM COMPLETED ----

SPATS-ESL adapts to each learner's needs by use of the Adaptive Item Selection (AIS) algorithm and by the action of the progressive curriculum. The progressive curriculum reduces the amount of time spent on constituent types as each type is mastered at Cumulative Level 4.

III. The Report Module and Client Feedback

Detailed reports of performance can be accessed by SPATS-ESL administrators. These reports include confusion matrices, information transmitted, IMS scores, and lists of confusions.

Clients are given feedback re their current performance in relation to their goals at the end of every run. They can follow their progress graphically as well.

IV. The Proctored Test Module

An ESL-Learner can schedule proctored tests with a SPATS-ESL administrator. In this way a student's performance can be certified in comparison to that of native speakers of English for any combination of Constituent Type and Level and on the Sentence task.

Results

Data obtained from 30 ESL learner's using SPATS-ESL will be described in a companion paper Friday at 3:15 p.

5pSWb1. Experience with computerized speech-perception training (SPATS-ESL) for speakers of other languages learning English. James D. Miller, Roy Sillings, Charles S. Watson, Communication Disorders Technology, Inc., 501 N. Morton St., Sta. 215, Bloomington, IN 47404, Isabelle Darcy, and Kathleen Bardovi-Harlig, Second Language Studies, Indiana Univ., Bloomington, IN 47405

It appears that most ESL-learners with a basic knowledge of English, pbt TOEFL scores near or well above 500, can approach the performance of native speakers of English after 15-35 hours of spaced practice on SPATS-ESL. This provides the ESL learner with the skills needed to learn more English through conversation with native speakers and to benefit from pronunciation instruction and self monitoring of speech productions.

References

1) Watson, C.S., Miller, J.D., Kewley-Port, D., Humes, L.E., and Wightman, F.L. (2008) Training Listeners to Identify the Sounds of Speech: I. A Review of past Studies. The Hearing Journal, 61(9), 26-31.
http://www.audiologyonline.com/theHearingJournal/pdfs/HJ2008_09_p26-31.pdf
2) Miller, J.D., Watson, C.S., Kistler, D.J, Preminger, J.E., and Wark, D. J. (2008) Training Listeners to Identify the Sounds of Speech: II. Using SPATS Software. The Hearing Journal, 61(10), 29-33.
http://www.audiologyonline.com/theHearingJournal/pdfs/HJ2008_10_p29-33.pdf
3) Miller, J.D., Watson, C.S., Kewley-Port, D., Sillings, R., Mills, W.B., and Burleson, D. F. (2008) SPATS: Speech Perception Assessment and Training System. Proceedings of Meetings on Acoustics, Vol. 2, 050005, 17 pp.
http://scitation.aip.org/getpdf/servlet/GetPDFServlet?filetype=pdf&id=PMARCW000002000001050005000001&idtype=cvips
4) Miller, J.D., Watson, C.S., Kistler, D. J., Wightman, F.L., and Preminger, J.L. (2008) Preliminary Evaluation of the Speech Perception assessment and Training system (SPATS) with Hearing-Aid and Cochlear-Implant Users. Proceedings of Meetings on Acoustics, Vol. 2, 050004, 9 pp.http://scitation.aip.org/getpdf/servlet/GetPDFServlet?filetype=pdf&id=PMARCW000002000001050004000001&idtype=cvips
5) Kingston, J. (2003) Learning Foreign Vowels. Language & Speech, 46(2-3), 295-349.

Supported by NIDCD SBIR Grant R44DC006338