acknowledgements

多様な英語能力の測定
--- Versant English Test, Versant Writing Test, Oxford Quick Placement Test などからの知見 ---

鈴木正紀（Pearson）
原田 康也（早稲田大学）

多様な英語力

- 大学入学者の英語力は多様である。
  - 初級から上級までが（同一クラス）に混在。
  - 知識と運用に乖離：A1-A2 vs B1-B2
  - 技能により習熟度が異なる
- 中級の学習者は多様である。
  - 初級の学習者は比較的均一に初級的である。
  - 上級の言語使用者は比較的均一に上級である。
  - 中級の特徴は知識・技能・運用が均一でないことである。
- 大学の英語授業カリキュラム・クラス編成
  - 「レベル分け」にどこまで意味があるのか？
  - 「スキル別」のクラス編成が適切であるのか？
- 4技能試験：CEFR or 高大接続
  - 合計点・平均点では意味がない、かも。

reality check: our methodology

- **Three tests, with Tera bytes of language data:**
  1. **Versant English Test (spoken English)**
  2. **Versant Writing Test (written English)**
  3. **Oxford Quick Placement Test**
- **Freshman students at Waseda University**
- **Administered Versant four times and OQPT three times in academic year 2015-2016**
  - **April, July, October*, December**

Versant English Test

- **Spoken English Test , 17 minutes**
- **Automated scoring with ASR**
- **Optimized for English learner speech patterns including Japanese learners of English**
- **Overall + four subscores**
  - **Sentence Mastery**
  - **Vocabulary**
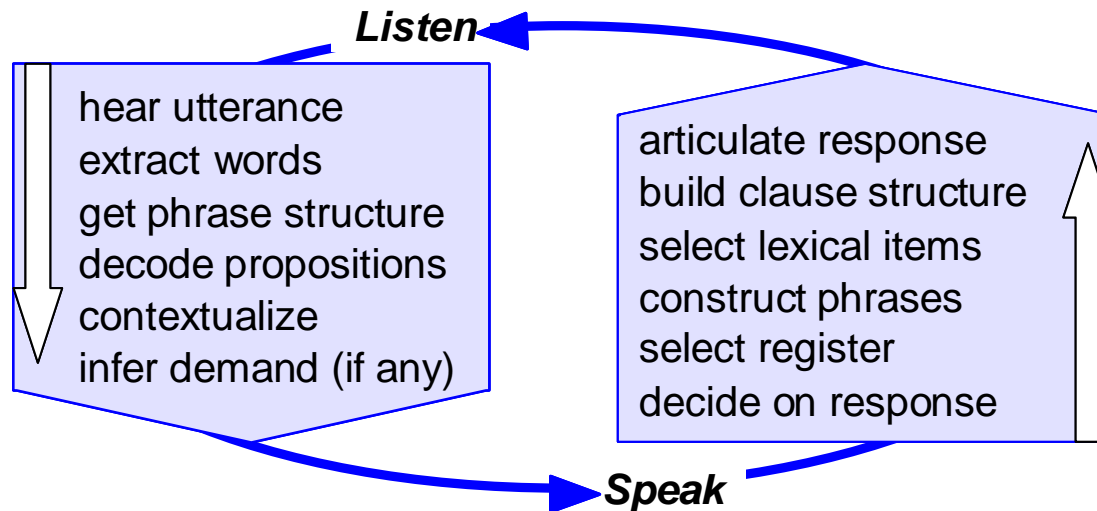 - **Fluency**
  - **Pronunciation**

Versant English Test

- **High frequency vocabulary in conversation (Most frequent 8,000 words in Switchboard Corpus）**
- **Relatively simple structures and expressions**
- **Items recorded at natural conversation pace by amateur native English speakers**
- **Focus on the ability to "listen, then speak" in real-time (=facility in spoken English)**
- **Psycholinguistic approach**

Test Construct

Facility in spoken English

The ability to understand spoken English on everyday topics and to respond appropriately and intelligibly in English at a native-like conversational pace.

Listen

hear utterance
extract words
get phrase structure
decode propositions
contextualize
infer demand (if any)

articulate response
build clause structure
select lexical items
construct phrases
select register
decide on response

Speak

Adapted from Levelt, 1989

Versant English Test

Part	Task
Part A	Reading （音読）
Part B	Sentence Repeat （復唱）
Part C	Short Answer Questions （質問）
Part D	Sentence Builds （文の構築）
Part E	Story Retelling （話の要約）
Part F	Open Questions （自由回答質問)*

*Open Questions are not automatically scored

Versant English Test

- **Test Reliability = 0.97 (split-half, test-retest)**
- **Machine-Human correlation = 0.97**
- **Concurrent validity evidence**

Test	N	Correlation
TOEFL iBT Speaking	130	0.75
IELTS Speaking	130	0.76
TOEIC Listening	171	0.71

VET - CEFR 対応表

VET	CEFR
20-25	>A1
26-35	A1
36-46	A2
47-57	B1
58-68	B2
69-78	C1
79-80	C2

From *Pearson (2003). Versant English Test Can-Do Guide.*

Relationship between Item Type and Score
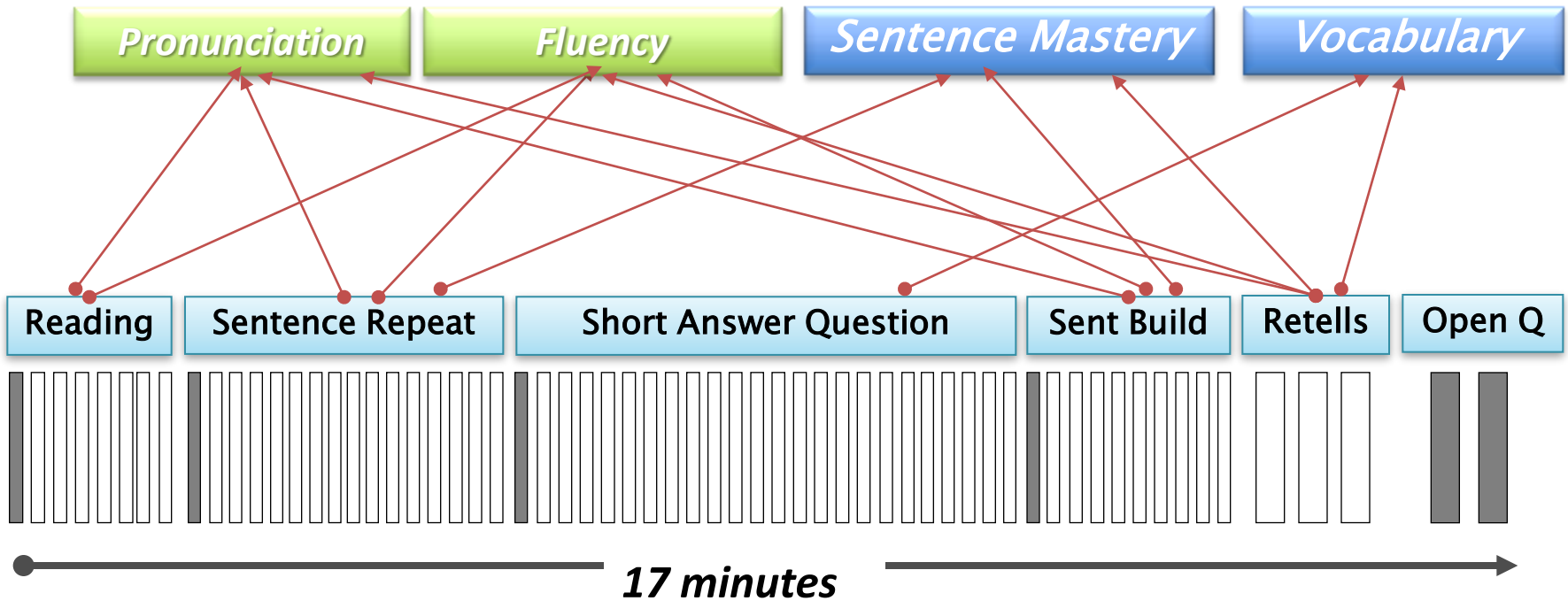
Automated Scoring System
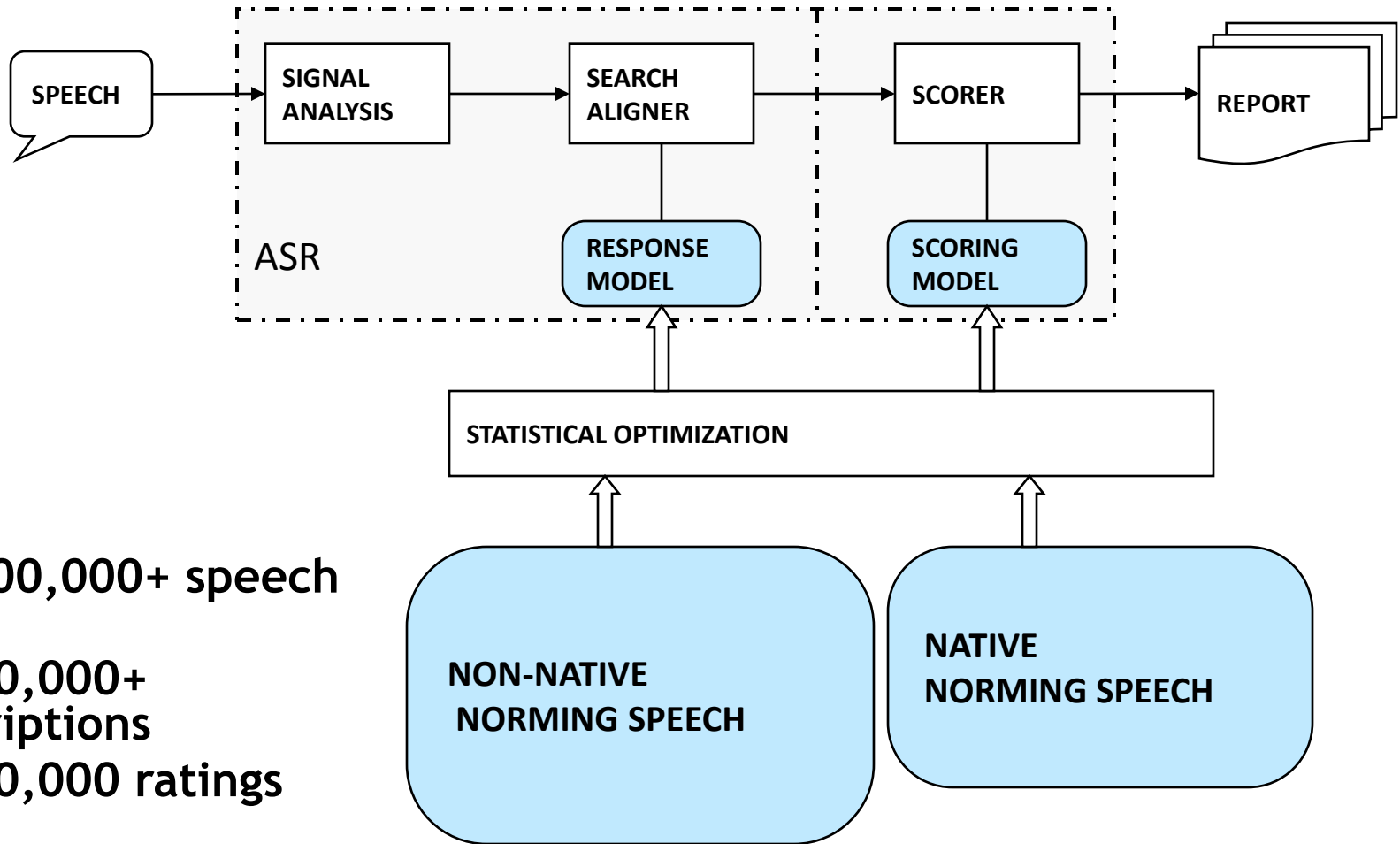
- **50,000,000+ speech files**
- **2,750,000+ transcriptions**
- **1,600,000 ratings**

Phone & Word Alignment
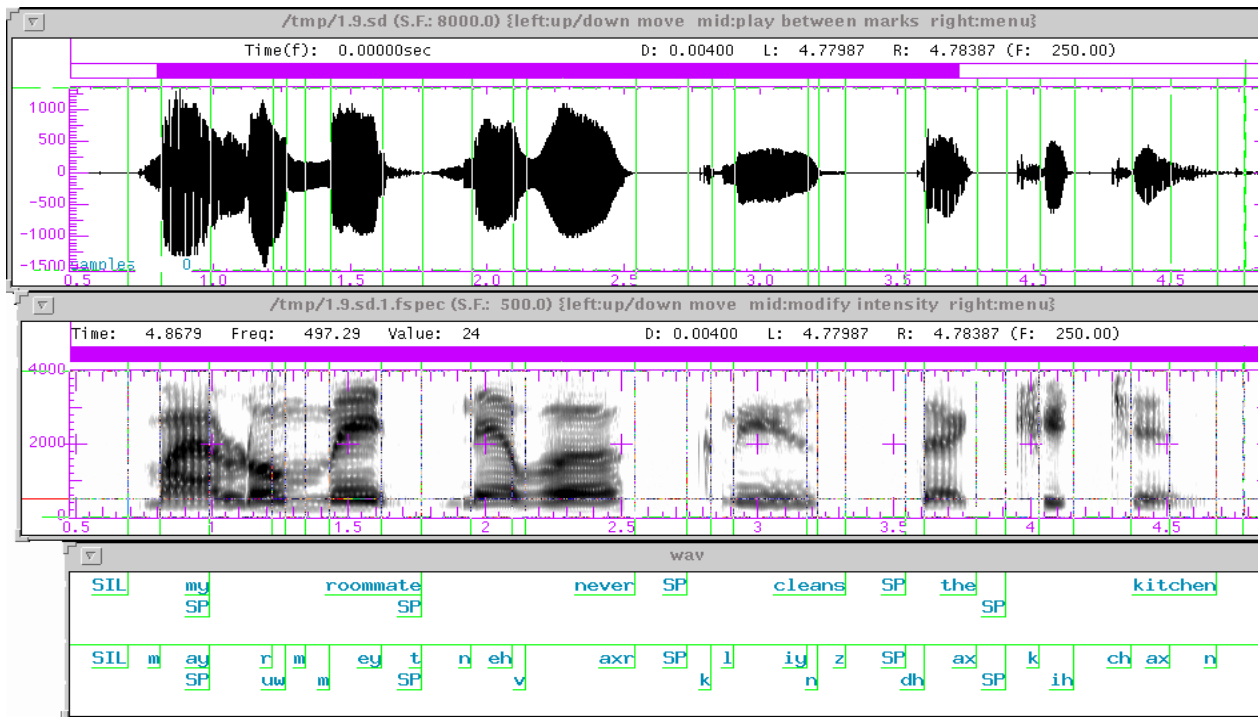
Simplified Response Network

Bill arrived at noon for his 11:30 appointment. How late was he?
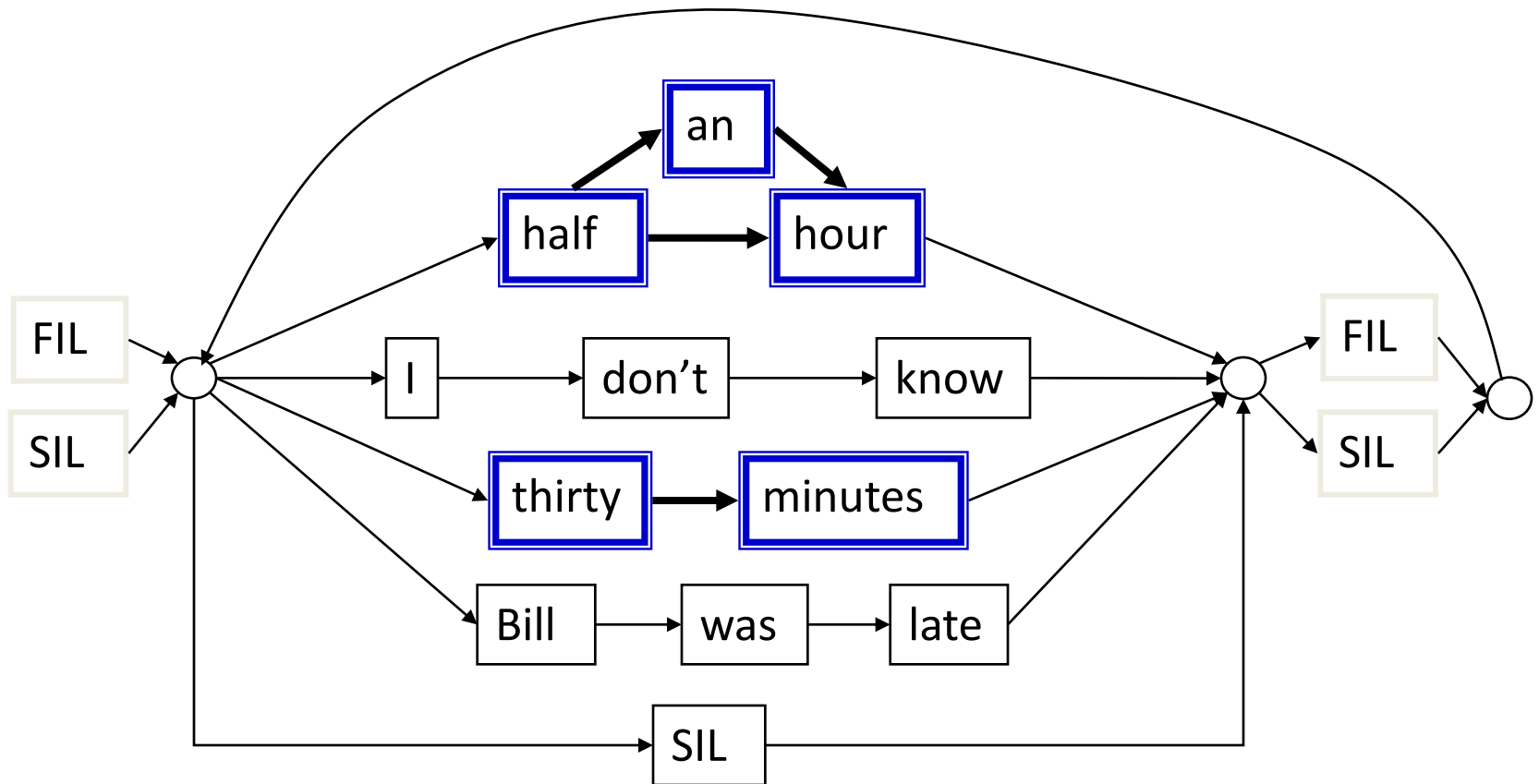
Versant Writing Test

- **Written English Test , ~40 minutes**
- **Automated scoring with Latent Semantic Analysis**
- **Overall + four subscores**
  - **Grammar**
  - **Vocabulary**
  - **Organization**
  - **Voice & Tone**
  - **Reading Comprehension**

Versant Writing Test

Part	Task
Part A	Typing
Part B	Sentence Completion
Part C	Dictation
Part D	Passage Reconstruction
Part E	E-Mail Writing

Versant Writing Test

- **Test Reliability = 0.98（split-half)**
- **Machine-Human correlation = 0.98**
- **Concurrent validity evidence**

| Test	N	Correlation
TOEIC	55	0.72

Oxford Quick Placement Test

- **Paper and pencil version (CD-ROM version includes listening)**
- **Focuses on Vocab, Grammar, Usage**
- **30 minutes, 60 multiple choice items (max 60 points)**
- **Recommended to be combined with speaking and writing test**
- **Concordance with CEFR levels**

The relation of the VET, VWT, and OQPT to the Common European Framework of Reference for Languages (CEFR)

VET	VWT	OQPT	CEFR level	CEFR level description
79-80	77-80	55-60	C2	Mastery (Upper advanced)
69-78	67-76	48-54	C1	Effective proficiency (Lower advanced)
58-68	54-66	40-47	B2	Vantage (Upper intermediate)
47-57	44-53	30-39	B1	Threshold (Lower intermediate)
36-46	30-43	18-29	A2	Waystage (Elementary)
26-35	20-29	10-17	A1	Breakthrough
20-25		0-9	<A1	(Beginner)

Versant Descriptive Statistics (Overall)

N = 53; Only freshman students who had both VET and VWT scores all four times

Versant English Test (VET)

VET	April	July	Sept	December
Mean	**38.89**	**41.47**	**40.57**	**42.42**
SD	6.77	6.45	6.18	5.54
Min	27	29	25	31
Max	55	62	56	60

Versant Writing Test (VWT)

VWT	April	July	Sept	December
Mean	**41.45**	**44.08**	**44.51**	**47.75**
SD	7.32	8.17	8.00	7.94
Min	20	20	22	22
Max	62	66	64	67

Mean Scores of VET, VWT Overall Scores
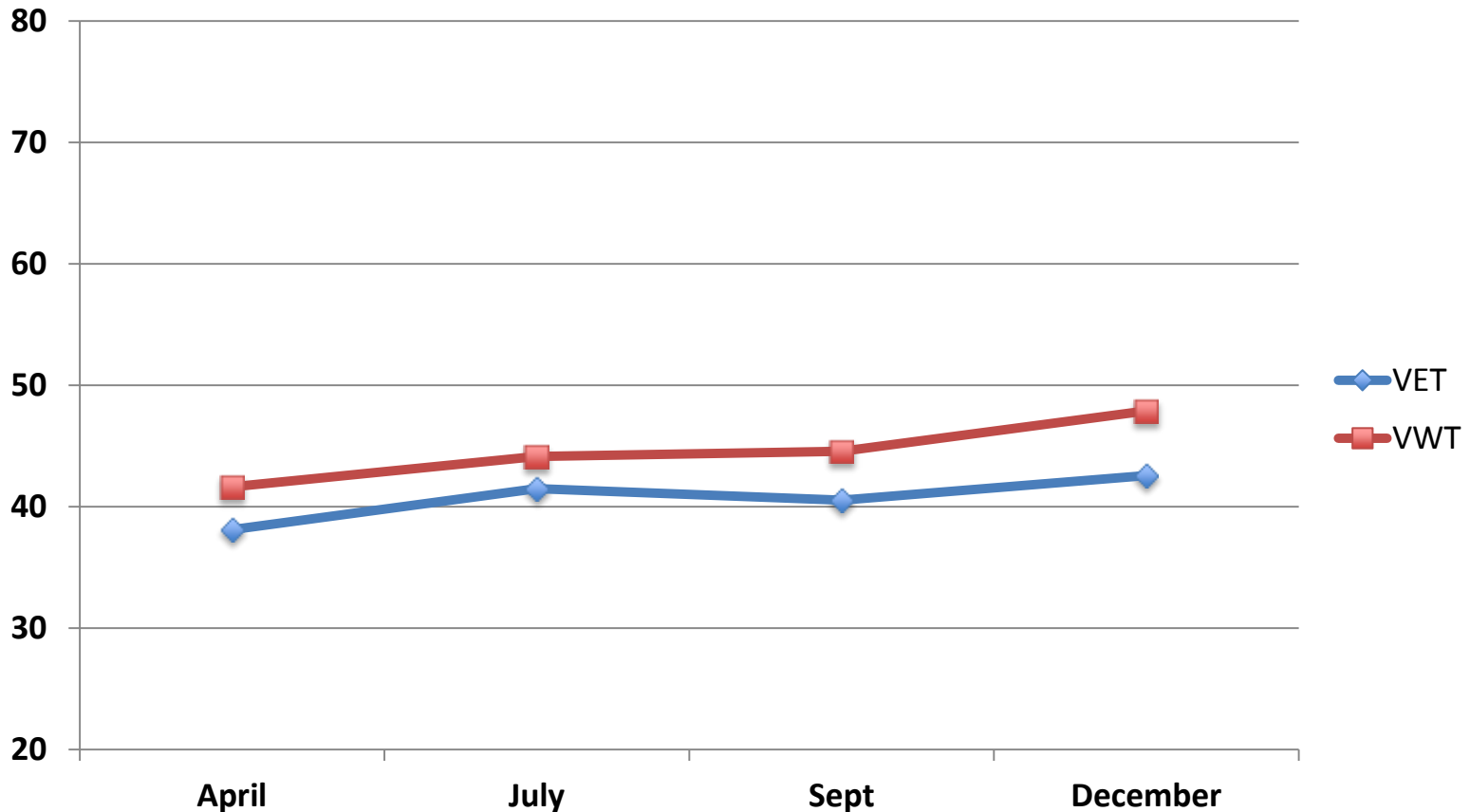
For both VET and VWT, the differences in mean scores between April and December are statistically significant (p<0.01)

Versant Descriptive Statistics (Subscores)

N = 53

Only freshman students who had both VET and VWT scores all four times

VET Subscore	April	July	Sept	December
Sent. Mastery	39.26	44.17	42.02	44.57
Vocabulary	37.34	40.68	41.00	43.68
Fluency	38.96	40.42	39.91	41.36
Pronunciation	39.23	39.79	38.66	39.62

VWT Subscore	April	July	Sept	December
Grammar	42.91	44.85	45.17	47.06
Vocabulary	40.47	42.77	42.08	45.30
Organization	40.45	43.91	45.98	50.13
Voice & Tone	40.75	44.04	44.89	50.25
Reading Comp.	42.19	45.42	46.45	50.42

Mapping VET scores to CEFR levels

N=53

CEFR	April	July	Sept	December
A1	19 (35.8%)	9 (17.0%)	10 (18.9%)	6 (11.3%)
A2	27 (50.9%)	36 (67.9%)	36 (67.9%)	35 (66.0%)
B1	7 (13.2%)	6 (11.3%)	7 (13.2%)	11 (20.8%)
B2	0	2 (3.8%)	0	1 (1.9%)
C1	0	0	0	0
C2	0	0	0	0

April-September: More than 80% is A1 and A2
December: Changed to A2 and B1

Mapping VWT scores to CEFR levels

N=53

CEFR	April	July	Sept	December
A1	3 (5.7%)	1 (1.9%)	2 (3.8%)	2 (3.8%)
A2	32 (60.4%)	22 (41.5%)	21 (39.6%)	10 (18.9%)
B1	16 (30.2%)	25 (47.2%)	23 (43.4%)	30 (56.6%)
B2	2 (3.8%)	5 (9.4%)	7 (13.2%)	10 (18.9%)
C1	0	0	0	1 (1.9%)
C2	0	0	0	0

April-Sept: More than 80% is A2 and B1
December: B1 is more than 55%; A1 and B1 are the same number

Oxford Quick Placement Test Descriptive Statistics

The same 53 freshman students.
April: One student did not take OXPT
December: 3 students did not take OQPT

	April (N=52)	July (N=53)	December (N=50)
Mean	**37.08**	**38.06**	**38.24**
SD	6.57	5.98	6.44
Min	20	25	21
Max	50	48	50

Mapping OQPT scores to CEFR levels

CEFR	April (N=52)	July (N=53)	December (N=50)
A1	0	0	0
A2	7 (13.5%)	4 (7.5%)	3 (6.0%)
B1	25 (48.1%)	25 (47.2%)	27 (54.0%)
B2	19 (36.5%)	21 (39.6%)	16 (32.0%)
C1	1 (1.9%)	3 (5.7%)	4 (8.0)
C2	0	0	0

April-December: B1 and B2 have more than 80% of the students

VET-VWT (April)

r=0.62

VET-VWT (December)

r=0.57

VET-OQPT (April)
OQPT Total vs VET Overall
r=0.25
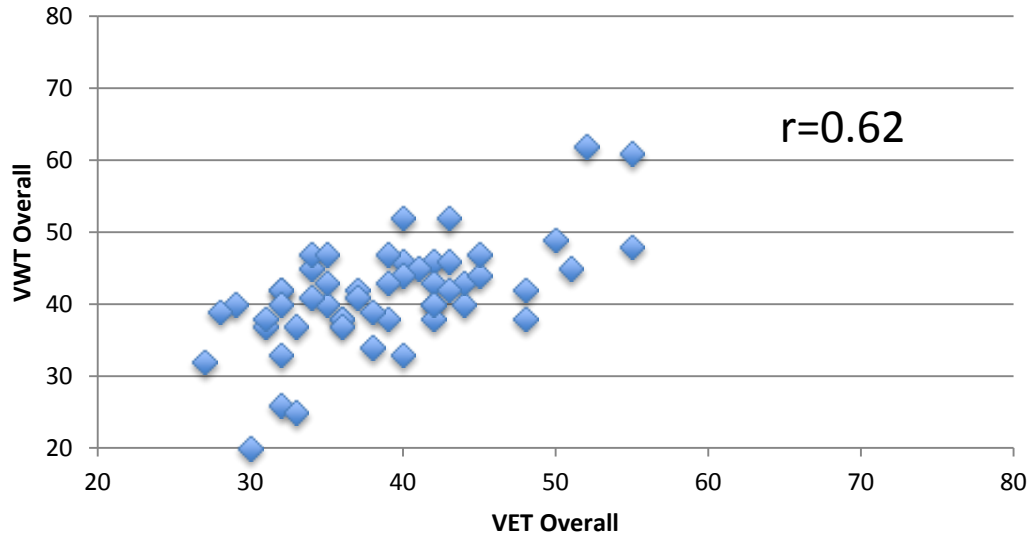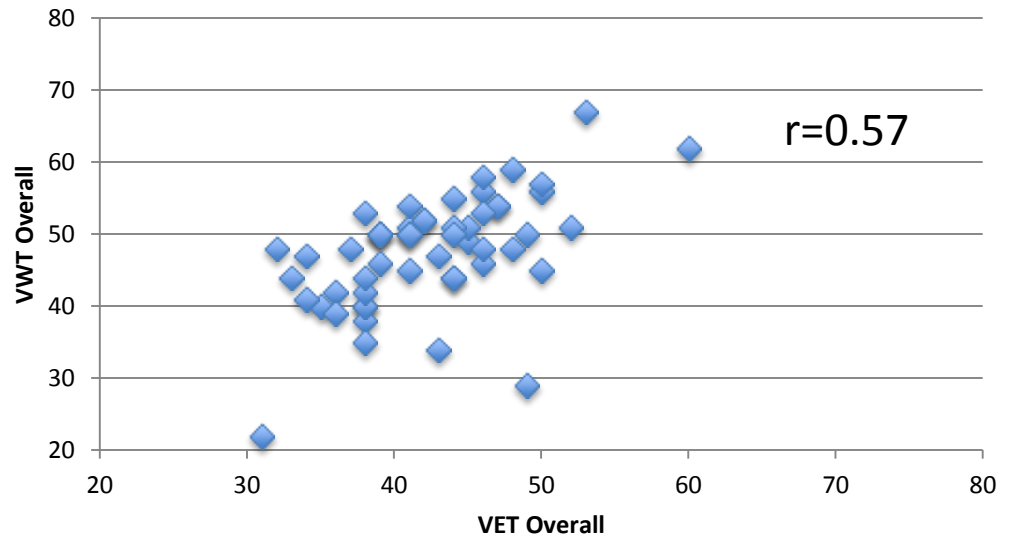
VET-OQPT (December)
OQPT Total vs VET Overall
r=0.49
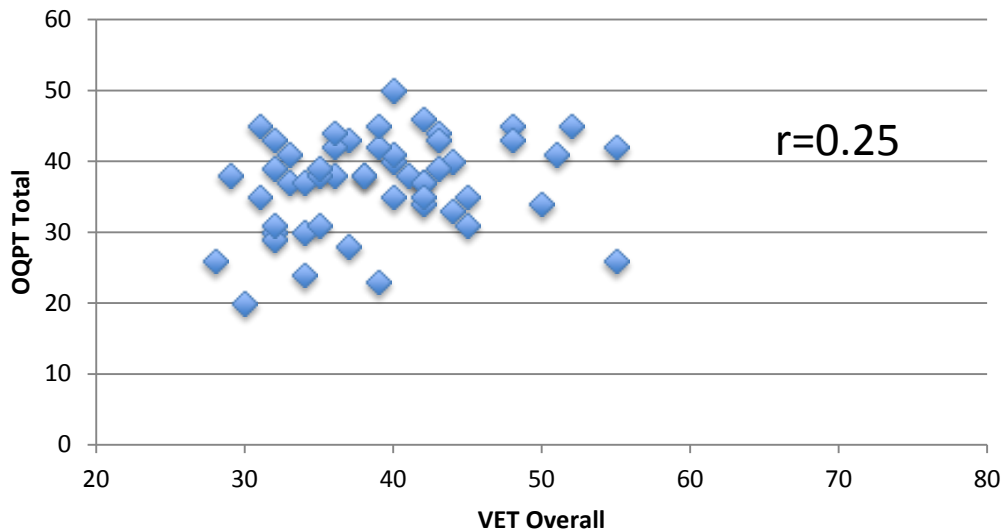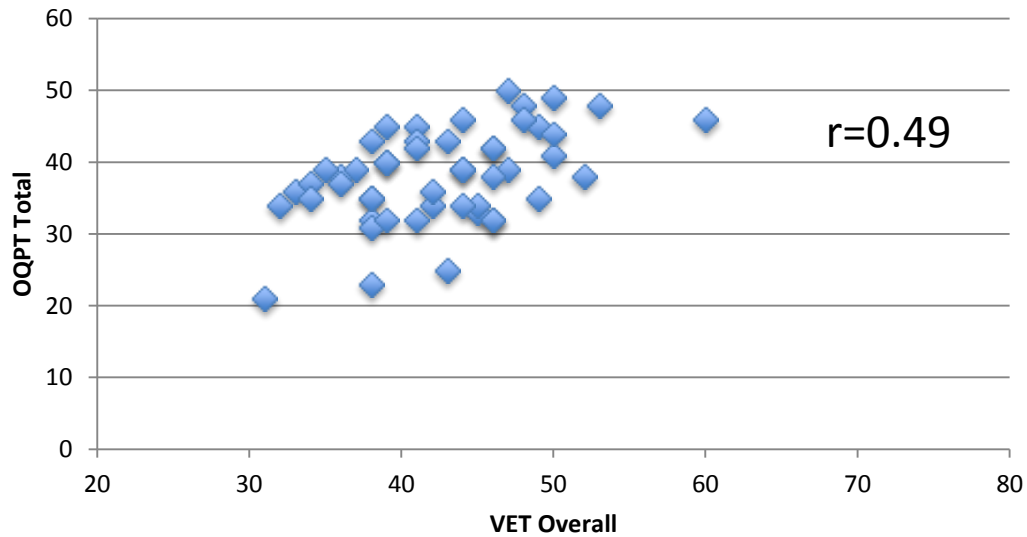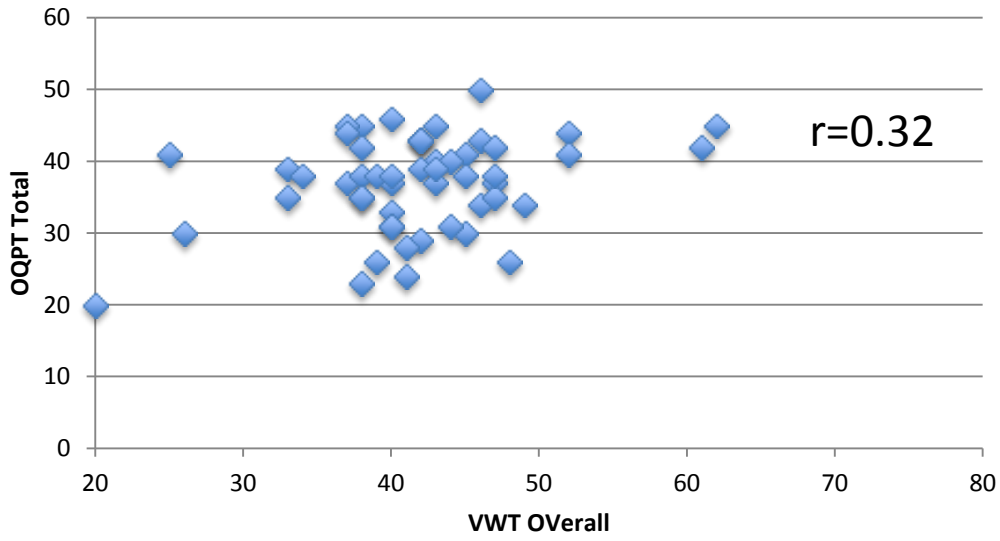
VWT-OQPT (April)

VWT-OQPT (December)
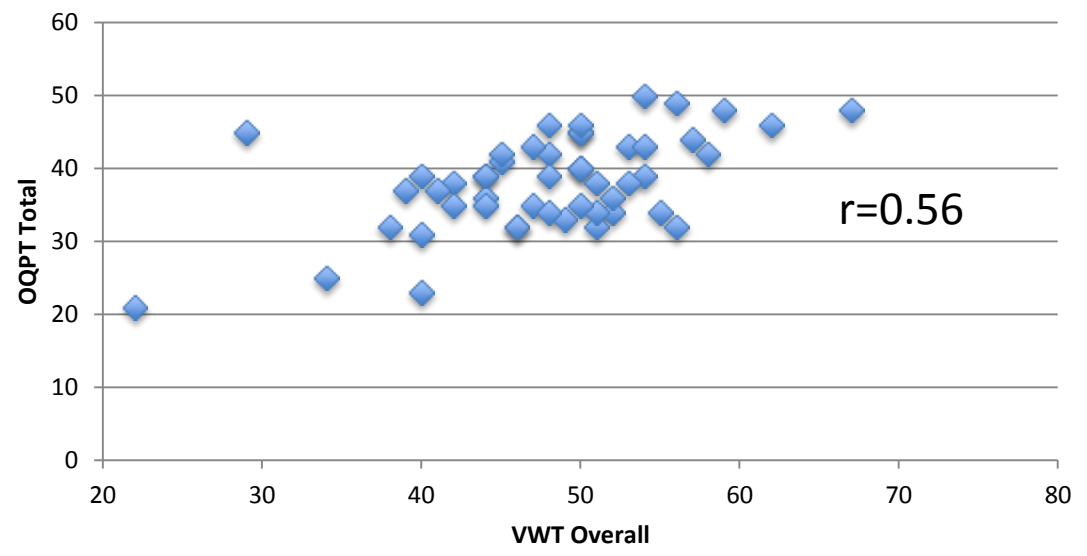
Summary

- **<u>VET</u>: scores improved between April and December. About 85% of the students were at A1 and A2 in April. In December, about 85% were at A2 and B1.**

- **<u>VWT</u>: scores improved between April and December. More than 80% of the students were at A2 and B1 in April. More students scored at B1 level in December.**

- **<u>OQPT</u>: scores did not change. 80% of the students were B1 and B2.**

Summary

- **Clear pattern exists for scores among the three tests: OQPT > VWT > VET**
- **Speaking is the weakest skill**
- **English classes focused on *using* English and its effect was more obvious in VET and VWT scores**
- **From knowledge to usage**