

New Approaches to C-test Development: Examples from Mandarin, Turkish, and Russian

Margaret E. Malone, Georgetown University, ACTFL

Yiran Xu, Seyma Toker, Derek Reagan, Georgetown University



GEORGETOWN UNIVERSITY



ACTFL

AMERICAN COUNCIL ON THE
TEACHING OF FOREIGN LANGUAGES

AGENDA

- C-test Background
- New Features in the AELRC C-test Projects
- Current AELRC C-test Projects
 - Mandarin
 - Turkish
 - Russian
- Conclusions & Future Directions

C-TESTS

- Similar to a cloze test in form
- Deletion method: 2-2-2 rule
 - Delete the second half of every second word, beginning from the second sentence (Raatz & Klein-Braley, 1981)
- 25 blanks for each text, 5 texts in total
- Quick and cost-effective
 - Open access; around 30mins
 - Placement tests (Mozgalina & Ryshina Pankova, 2015; Norris, 2006; Wilmes, 2007)
- Good indication of global or integrated language proficiency (e.g., Eckes & Grotjahn, 2006; Babaii & Ansari, 2001; Babaii & Jalali Moghaddam, 2006)

Learning to write

I was four when I started to learn to write. My grandfather started to teach me before I went to school. I remember th I always fo the cap letters mu easier t write th the sm letters. I reme that on we sta writing i school w were n allowed t use pe , we h to u pencils un we bec really go at wri . I can wr a few chara in Chi now, b not very many. As far as my normal writing is concerned, I have two styles of handwriting: one for other people and one for myself. Most people think that when I am really trying hard, my writing is easy to read and looks quite good. I think it looks quite impressive too.

Correct

Answers

Reset

PREVIOUS C-TEST RESEARCH

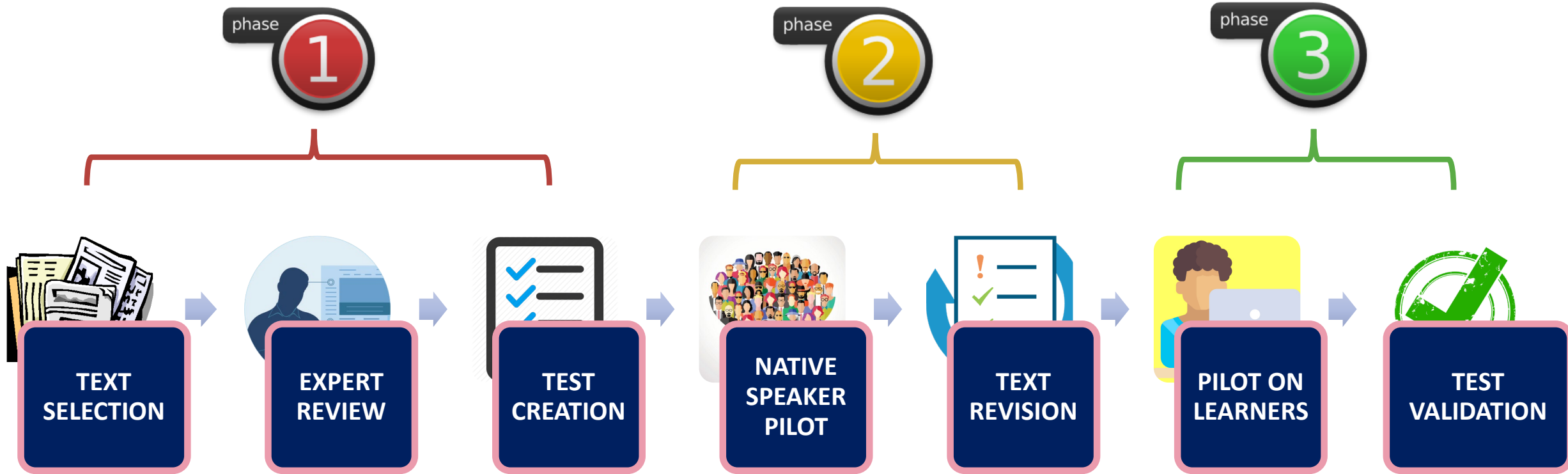
Letter-based scripts

- **English** (e.g., Baghaei, 2008; Babaii & Ansary, 2001; Harsch & Hartig, 2015; Jafarpur, 1995)
- **French** (Grotjahn & Stemmer, 1985; Reichert, Keller & Martin, 2010; Ward, 1987)
- **German** (Eckes & Grotjahn, 2006; Mozgalina & Ryshina-Pankova, 2015; Norris, 2006; Wilmes, 2007)
- **Portuguese** (Maimone, 2018)
- **Turkish** (Demiralp, 2018)

Character-based scripts

- **Korean** (Lee-Ellis, 2009; Son, Kim, Cho & Davis, 2018)
- **Japanese** (Sasayama, 2018)
- **Bangla** (McKay & Abedin, 2018)
- **Mandarin** (Arras & Grotjahn, 1994; Zhang, 1985)

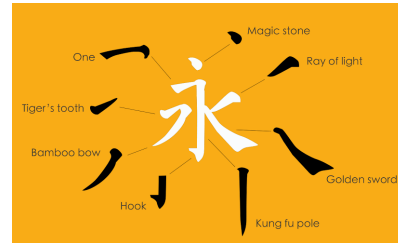
C-TEST DEVELOPMENT & VALIDATION PROCEDURE



NEW FEATURES IN OUR C-TEST PROJECTS



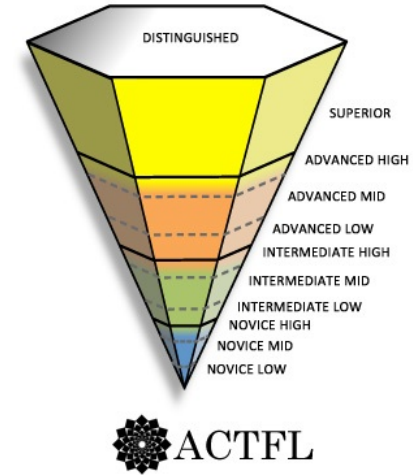
Think-aloud
(Mandarin, Turkish, & Russian)



Two Deletion Methods:
Word-based & Stroke-based
(Mandarin)

听 → 聽

Two versions available:
Simplified & Traditional
(Mandarin)



Text selection based on
ACTFL Guidelines
(Turkish)



Validation with ACTFL tests
(Mandarin)

THE C-TEST PROJECTS AT AELRC

- Japanese (Sasayama, 2018)
- Korean (Son, Kim, Cho, & Davis, 2018)
- Bangla (McKay & Abedin, 2018)
- Mandarin (Xu & Malone, in preparation)
- Turkish, English, and Russian (ongoing projects)

Doctoral Dissertations

- Korean C-test for heritage language learners (Son, 2018)
- A meta-analysis of C-test studies (McKay, 2019)

The Mandarin C-test Project

A sample text in simplified characters



Mandarin C-test (Xu & Margaret, 2019)

No. 1.

我的名字叫马克，我来自美国纽约，我来中国已经三年了。我现 (1)_____ 是北京大 (2)_____ 三年 (3)_____ 的学生，我在这 (4)_____ 学习中 (5)_____ 和国 (6)_____ 关系。我喜 (7)_____ 旅行，美 (8)_____，运动和电 (9)_____。周末的时 (10)_____ 我喜欢跟我的中国朋 (11)_____ 一起打打篮 (12)_____，看看电影，或 (13)_____ 在北京附 (14)_____ 玩一玩，品 (15)_____ 地方的小 (16)_____。我也很喜欢跟朋友去酒 (17)_____ 聊天喝 (18)_____，认识更多新的朋友。我很想 (19)_____ 我的家 (20)_____，我每 (21)_____ 都会跟他们打电 (22)_____。我的爸爸妈妈告 (23)_____ 我他们想来中国看我，我很高 (24)_____。到时候，我会带他们去爬长城，看故宫，吃一些北京的特 (25)_____ 小吃。我已经等不及要见到他们了！

A sample text in traditional characters



Mandarin C-test (Xu & Margaret, 2019)

No. 1.

我的名字叫馬克，我來自美國紐約，我來中國已經三年了。我現 (1)_____ 是北京大 (2)_____ 三年 (3)_____ 的學生，我在這 (4)_____ 學習中 (5)_____ 和國 (6)_____ 關係。我喜 (7)_____ 旅行，美 (8)_____，運動和電 (9)_____。週末的時 (10)_____ 我喜歡跟我的中國朋 (11)_____ 一起打打籃 (12)_____，看看電影，或 (13)_____ 在北京附 (14)_____ 玩一玩，品 (15)_____ 地方的小 (16)_____。我也很喜歡跟朋友去酒 (17)_____ 聊天喝 (18)_____，認識更多新的朋友。我很想 (19)_____ 我的家 (20)_____，我每 (21)_____ 都會跟他們打電 (22)_____。我的爸爸媽媽告 (23)_____ 我他們想來中國看我，我很高 (24)_____。到時候，我會帶他們去爬長城，看故宮，吃一些北京的特 (25)_____ 小吃。我已經等不及要見到他們了！

MANDARIN LANGUAGE

- Non-alphabetic; based on a logographic system
- Constructed horizontally from left to right
- No spacing between words and characters
- What is a word in Chinese?
 - A minimal linguistic form that has meaning, can occur independently in speaking and writing (Fu, 1985; Zhu, 1982) and has translation equivalent in other languages (Chao, 1968) → (therefore, particles like “ba” “le” “ma” do not count as words)
 - It may include 1 character 我 (I); or 2 characters 旅行 (travel) or more

PHASE 1 TEXT SELECTION & EXPERT REVIEW

TEXT SELECTION (20 texts)

- A variety of familiar genres
- All from authentic sources (included texts from textbook for school-age children for lower-levels)
- Generic topics
- Self-contained
- Texts were selected based on the descriptors of HSK (Hanyu Shuiping Koashi; translated as the Chinese Proficiency Test) administered by Hanban



NEWSPAPER

- Text 1, 6, 8, 19
- Entertainment, life style, social network, culture



MAGZINE

- Text 13, 14, 17, 18
- Health, anthropology, travel, communication



POPULAR MEDIA

- Text 2, 3, 5, 7
- Commentary, travel descriptions, product description, customer review



BLOGS, LETTERS, EMAILS

- Text 4, 15, 16, 20
- Education, politics, family, cover letter



TEXTBOOK FOR SCHOOL-AGE CHILDREN

- Text 9, 10, 11, 12
- Attitude, philosophy, animal, travel

EXPERT REVIEW



- 25 Chinese instructors with 3+ years teaching experience (average: more than 10 years)
- Teaching level: university (21), secondary (1), elementary (1), mixed (2)
- Recruited from more than 100 U.S. universities
- Education: Master's (12), PhD (13)

Survey

- Rate the difficulty level of 20 texts (1-5, 1 being very easy, 5 being very difficult)
- Provide qualitative feedback

- Eliminated 5 texts (out of 20)
 - Technical
 - Culture-specific
 - Required background knowledge
- Added two lower-level texts (manipulated texts)
 - Self-introduction
 - Narrative of daily activities

PHASE 2 TEST CREATION & NATIVE SPEAKER PILOT

TWO DELETION METHODS

Word-based deletion: delete the second half of every other word from the second sentence

(15) 我的名字叫马克，我来自美国纽约，我来中国已经三年了。我现_____是北京大_____三年_____的学生，我在这_____学习中_____和国_____关系。我喜_____旅行，美_____，运动和电_____。周末的时_____我喜_____跟我

Stroke-based deletion: delete the second half (by the order of the stroke) of every other character from the second sentence

我的名字叫马克，我来自美国纽约，我来中国已经三年了。我 王 在 旦 北京大学三 个 级的学 个，我在 之 里学 丁 中 二 和 日 际 兰 系。我喜欢 旅 行， 兰 食， 云 动和 口 影。 冂

Deletion Rules & Scoring Principles

- ICTCLAS tokenization tool
(<http://ictclas.nlpir.org/nlpir/>)
- Manual checking

Scored dichotomously (0 or 1)

Score range for each text: 0 -25

Accept alternative answers

Accept both simplified and traditional characters

Accept orthographic errors if they do not interfere with meaning



NATIVE SPEAKER PILOT

GROUP 1: Word-based

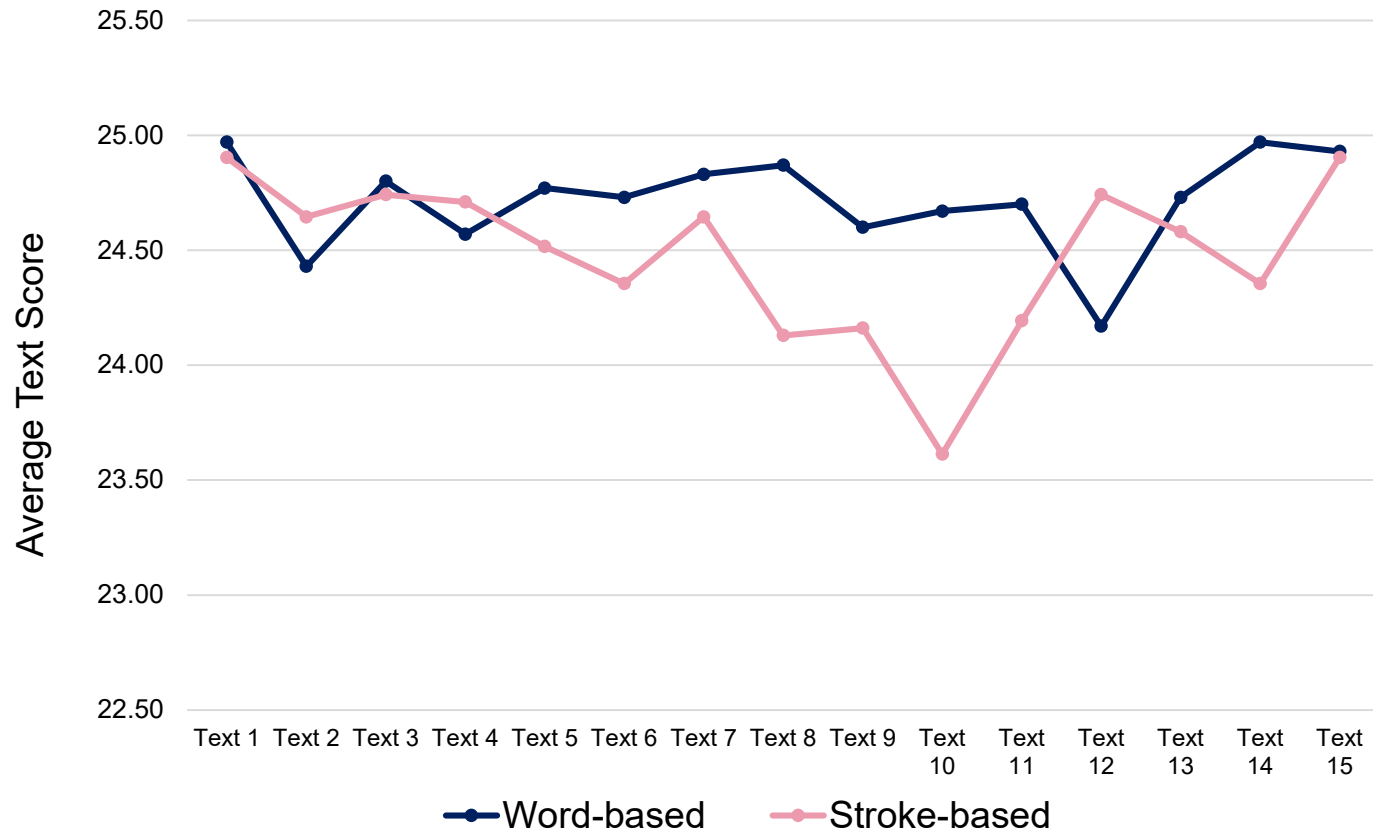
- 30 Mandarin native speakers living in the U.S. and China
- Age: $M = 21.6$; $SD = 2.89$
- Think-aloud ($N=4$)
- Undergraduate and graduate students
- Provided feedback on text difficulty



GROUP 2: Stroke-based

- 31 Mandarin native speakers living in the U.S. and China
- Age: $M = 25.56$; $SD = 3.34$
- Think-aloud ($N=4$)
- Undergraduate and graduate students
- Provided feedback on text difficulty

NS PILOT RESULTS

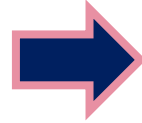


- Range: 23.6 – 25
- Average accuracy rate:
 - Word-based: 98.87%
 - Stroke-based: 97.92%
- More variations in the stroke-based version
- Think-aloud & qualitative feedback:
 - The stroke based is more challenging because it does not allow for alternative answers
 - Sometimes, no contextual information is needed to complete the character for the stroke-based version

PHASE 3 TEXT REVISION & LEARNER PILOT

TEXT REVISION

- Word-based deletion
- Deleted another 5 texts (out of 15)
 - Challenging for NSs
 - Translated articles
- Revised the remaining texts (based on feedback & think-aloud data)

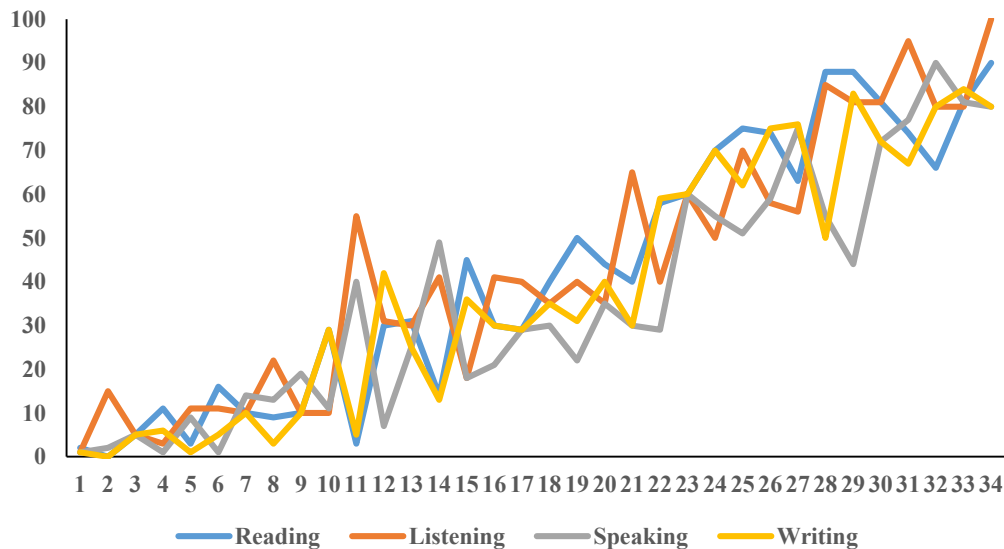


TEXT TOPICS

TEXT 1: Manipulated text: Self-introduction
TEXT 2: Manipulated text: Visiting a friend in hospital
TEXT 3: Review: A review of a school-age book
TEXT 4: Cover letter: Job application
TEXT 5: Customer Reviews: Bubble tea
TEXT 6: Opinion: Livestreaming industry in China
TEXT 7: Opinion: The philosophy of growing flowers
TEXT 8: Description: Traveling in Tibet
TEXT 9: Description: Avocado
TEXT 10: Comments: Museums in China

LEARNER PARTICIPANTS

- Computer-based (or Paper-based test upon request) (40-50mins) + background questionnaire (5mins)



Learners of Chinese (N = 34)

- Age: Mean 22.9 (*SD* 5.11)
- Gender: 12 M; 22 F
- Recruited from 3 major universities in the U.S.
- 19/34 (Currently enrolled in a Chinese program)
- Think-aloud (N = 21)
- ACTFL OPIc & RPT (N = 24)

Self-reported proficiency (on a scale of 0 to 100)

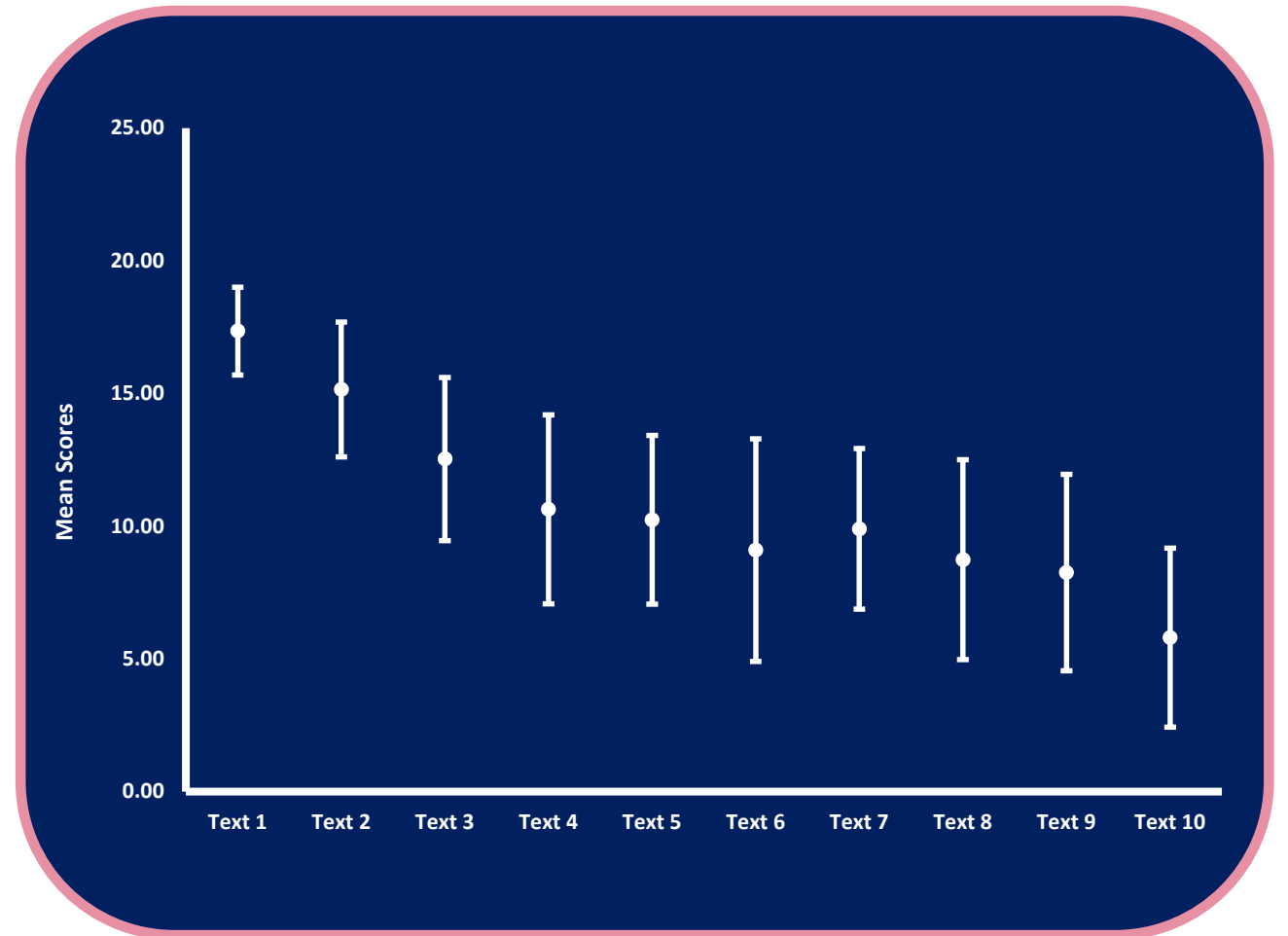
- Reading (M = 41.74, SD = 29.78, Range 0-91)
- Writing (M = 38.35, SD = 28.91, Range 0-80)
- Listening (M = 43.09, SD = 29.07, Range 1-100)
- Speaking (M = 35.59, SD = 26.97, Range 1-80)

Average self assessment in four levels

- Level 1: 10 (0-25)
- Level 2: 12 (25-50)
- Level 3: 7 (50-75)
- Level 4: 5 (75-100)

DESCRIPTIVE STATISTICS

- Possible range: 0-25
- Actual range: 0-25
- Mean scores: 5.8 – 17.35
- Cronbach's alpha: 0.98



IRT ANALYSIS

Super-items:scale 0-25
Rating Scale Model

Summary fit statistics within range

- Person separation index: 7.73
- Separation reliability: 0.98
- Real RMSE: 0.18
- Item separation index: 7.73
- Separation reliability: 0.99
- Real RMSE: 0.09

```
<more proficient>|<more difficult>
5              4              +
              |
              X |
              |
3              +
              T|
              |
              XX|
2              X +
              X |T
              |
              X S| T10
1              X +
              |S T8
              |T9
              |T6 T7
0              X +M T4 T5
              XXXX|
              XXX M| T3
              XXX |S
-1              X + T2
              XX |
              XXX |T1
              XX |T
-2              XX S+
              |
              |
-3              +
              |
              XX |
              |
              X T|
-4              +
<less proficient>|<less difficult>
```

SELECTED FIVE TEXTS RESULTS

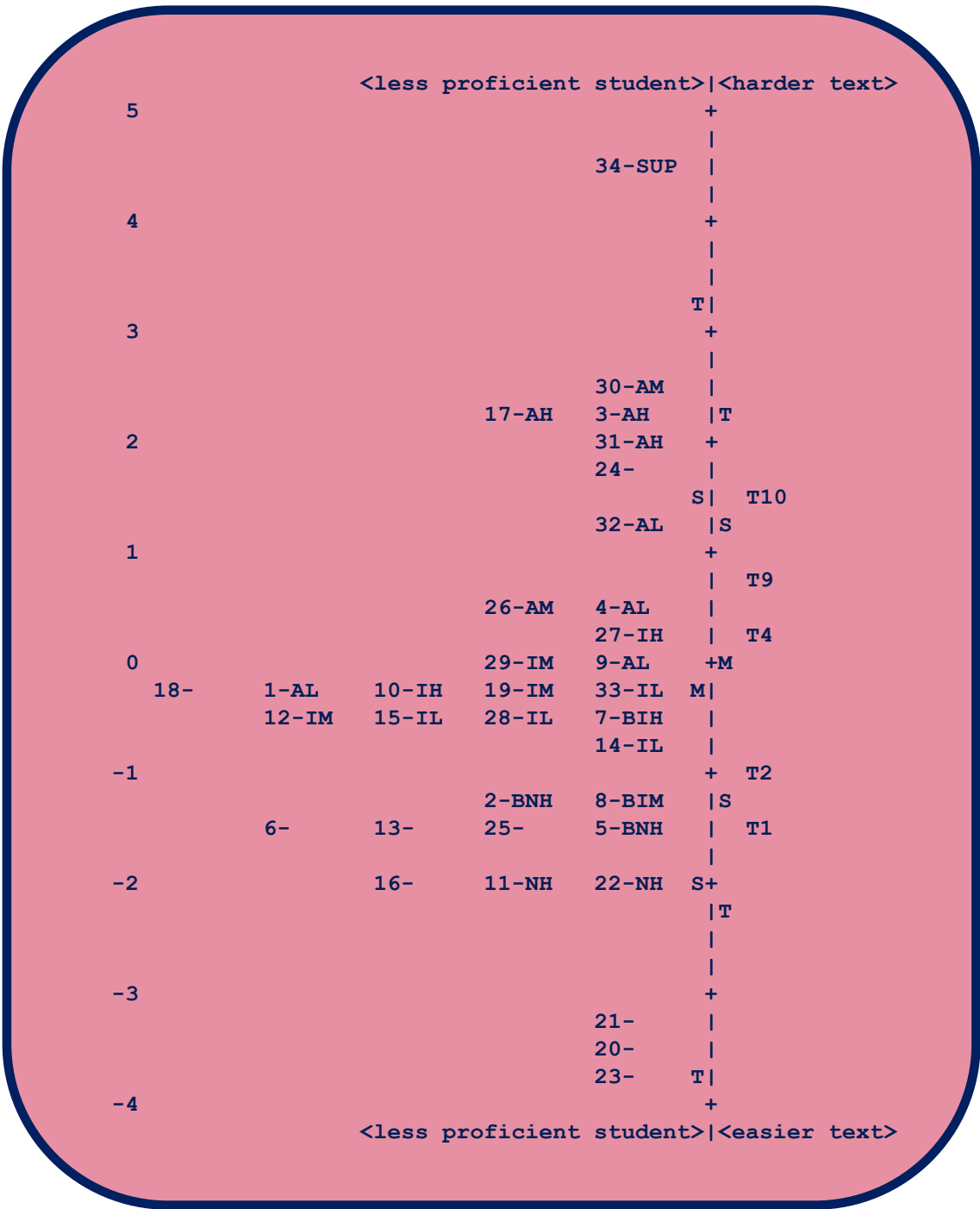
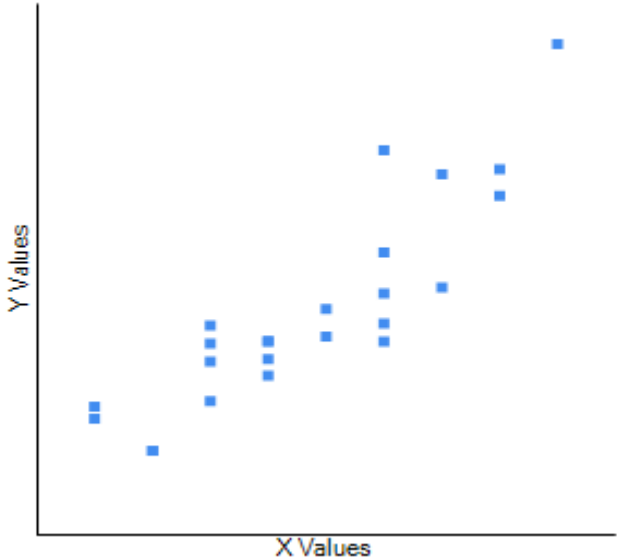
Summary fit statistics within range

- Person separation index: 6.52
- Separation reliability: 0.94
- Real RMSE: 0.32
- Item separation index: 13.14
- Separation reliability: 0.99
- Real RMSE: 0.10

```
<less proficient|<less difficult>
6      +
      |
      X |
5      +
      |
4      +
      T|
      X |
3      X +
      X |T
      X |
2      X + T10
      S|
      X |S
1      +
      XX | T9
      X | T4
0      XX +M
      XXXXXXX M|
      XX |
-1     X + T2
      X |S
      XXXXX | T1
-2     +
      XXX S|
      |T
-3     +
      |
      X |
-4     X +
      X |
      T|
-5     +
      <less>|<freq>
<less proficient>|<less difficult>
```


CORRELATIONS: C-TEST & ACTFL RPT RATINGS

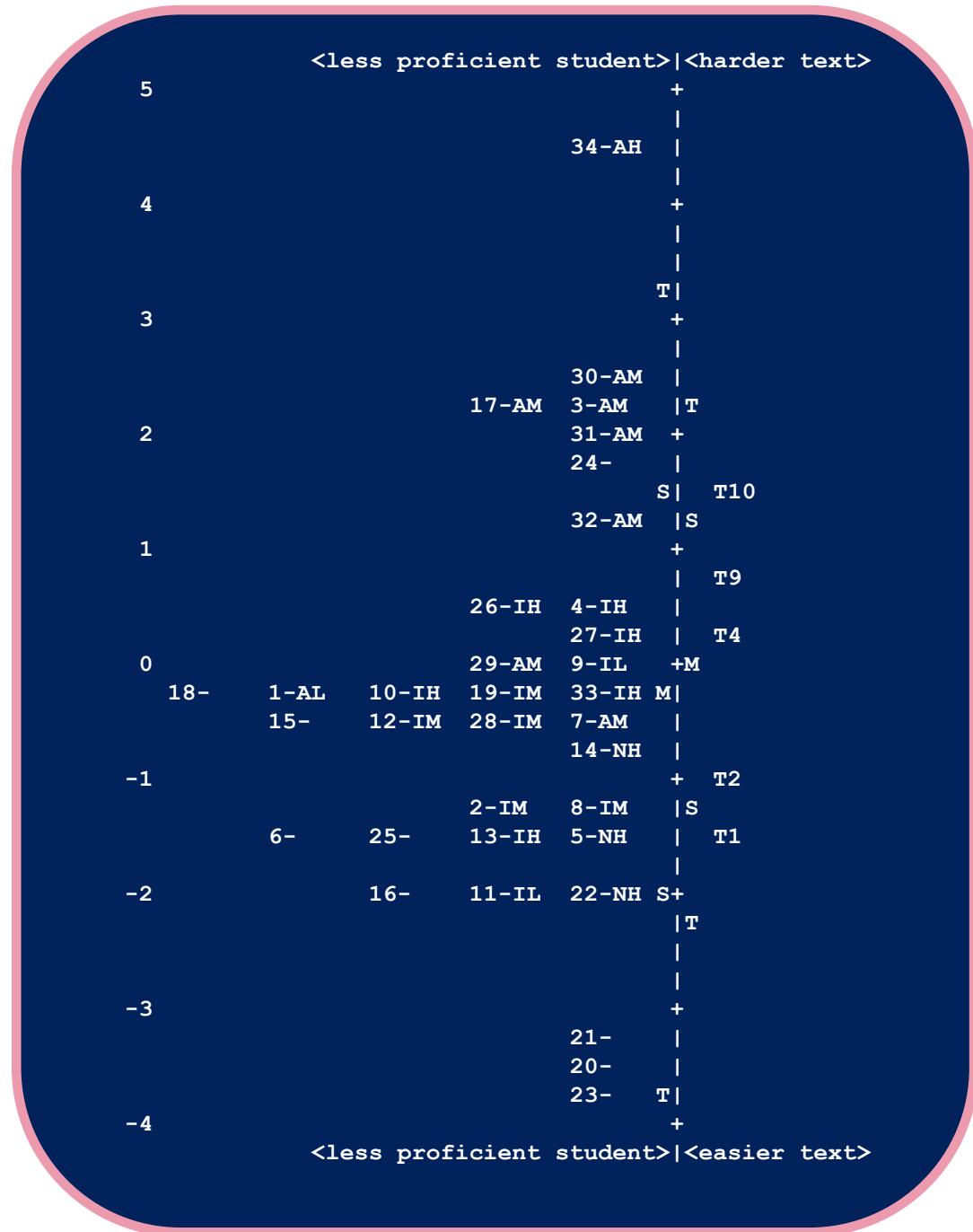
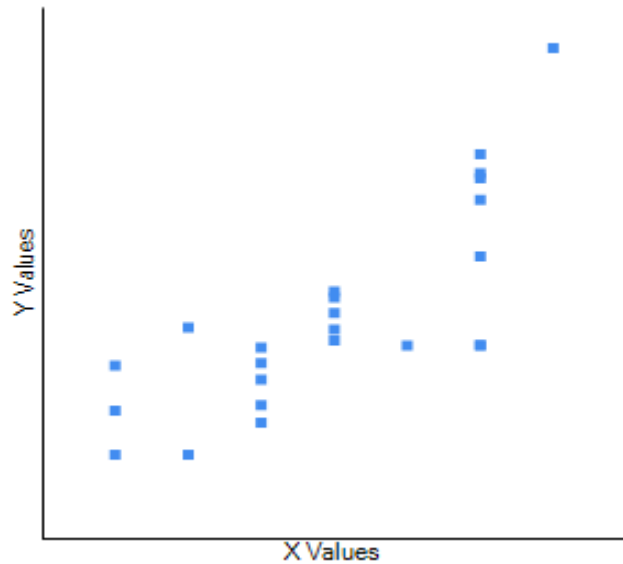
Reading Proficiency Test & C-test
logit scores on 5 selected texts:
 $r= 0.91$ ($p < 0.01$, $n = 24$)



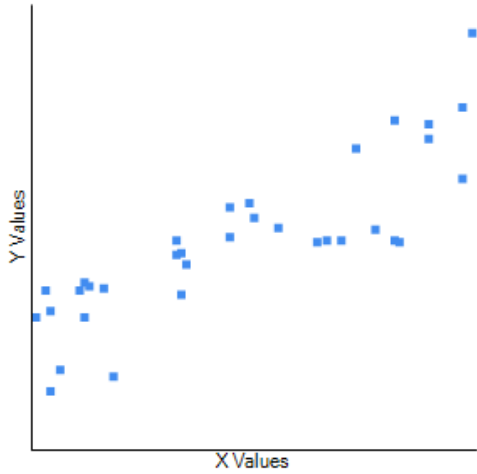
CORRELATIONS: C-TEST & ACTFL OPIc RATINGS

OPlc & C-test logit scores on 5 selected texts:

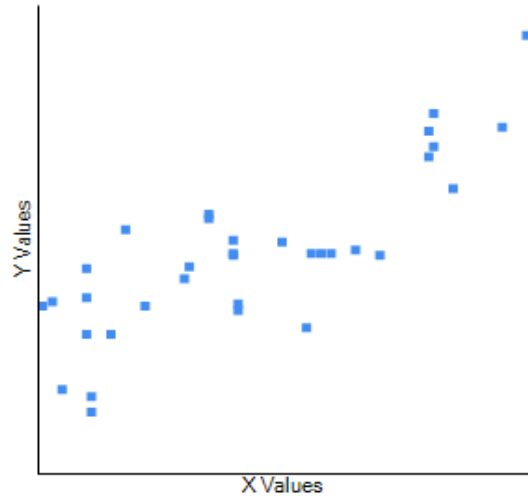
$$r = 0.81 (p < 0.01, n = 24)$$



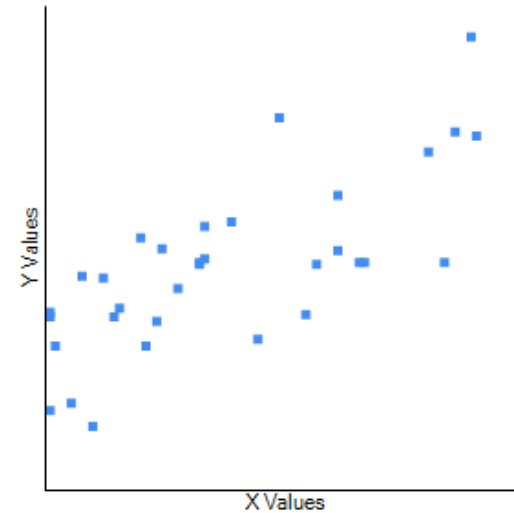
CORRELATIONS: C-TEST & SELF-ASSESSMENT



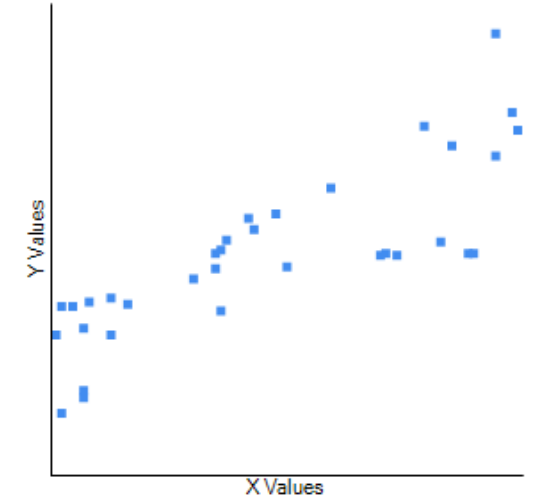
Reading ($r = 0.85, p < 0.01$)



Listening ($r = 0.85, p < 0.01$)



Speaking ($r = 0.75, p < 0.01$)



Writing ($r = 0.75, p < 0.01$)

DISCUSSION

- A reliable tool for SLA research purposes
 - Quick (30 mins) and free (upon request)
 - High reliability indices
 - Distribute learners into different major levels
 - Measure global language proficiency
 - Highly correlated with ACTFL RPT and OPIc ratings
 - Highly correlated with self-assessed proficiencies in four skills
- However...
 - Since Chinese is a logographic-based language, it is challenging for learners with limited literacy skills to meet basic demands of the test
 - Correlate more with receptive than productive language skills

The Turkish C-test Project

THE TURKISH C-TEST PROJECT

Goal

- Revise the previous Turkish C-test (Demiralp, 2018), which was designed according to ILR reading scale, to align with the ACTFL proficiency guidelines

New Features

- ACTFL guidelines for text selection
- Holistic rubric for expert review
- Think-aloud protocol in interviews with experts

PHASE 1 TEXT SELECTION & EXPERT REVIEW

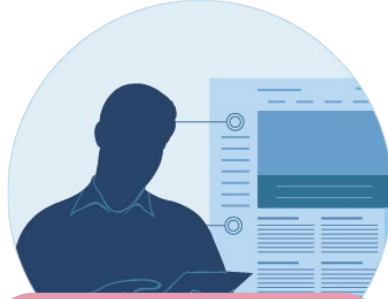
phase

1



TEXT SELECTION

- 11 validated texts from Demiralp (2018)



EXPERT REVIEW

- Survey and interview with 10-15 Turkish language instructors

Text 1



Burası benim mahallem. Benim evim ana caddede. Evimin karşısında bir lokanta var. Lokantanın servisi güzel, ama fiyatları biraz yüksek. İki sokak ileride bir bakkal var, ama büyük değil. Arka sokakta büyük bir süpermarket var. Genellikle orada alışveriş yaparım. Yan sokakta küçük bir sinema var. Film izlemek için güzel bir yer.

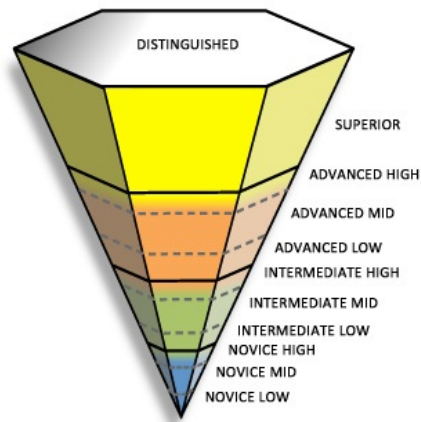
14. In your opinion, what is the proficiency level of this text?

☐ Novice

☐ Intermediate

☐ Advanced

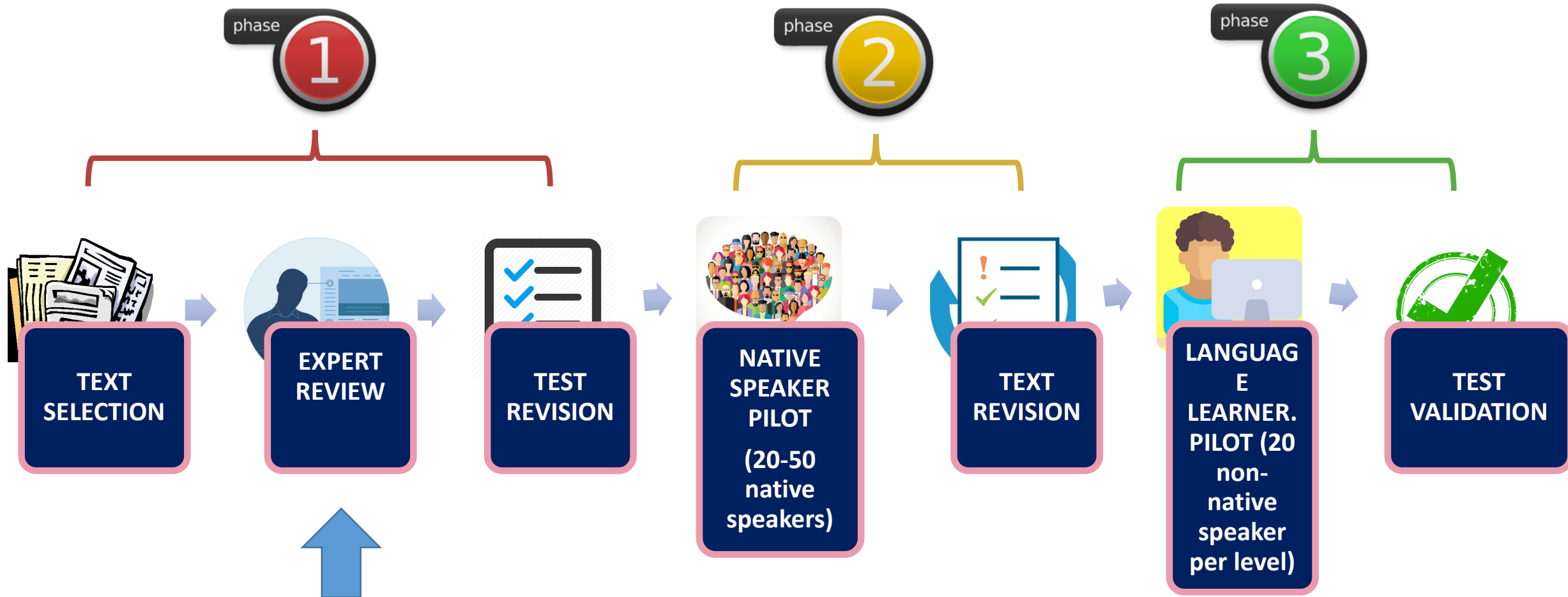
PHASE 1 TEXT SELECTION & EXPERT REVIEW



15. Now, read the items under each level and check all that apply.

Novice	Intermediate	Advanced
<p><input type="checkbox"/> The text includes key words, cognates, and formulaic phrases that are highly contextualized.</p> <p><input type="checkbox"/> The information in the text is highly predictable.</p> <p><input type="checkbox"/> The text is about a highly familiar topic that is relevant to daily life</p> <p><input type="checkbox"/> The readers can easily draw on their background knowledge to understand the text.</p> <p><input type="checkbox"/> The text includes words with letters unique to Turkish alphabet (e.g. "ç, ı, ğ, ö, ş, ü")</p>	<p><input type="checkbox"/> The text is primarily organized in individual sentences and strings of sentences containing predominantly high frequency vocabulary.</p> <p><input type="checkbox"/> The information in the text is conveyed in simple, predictable, loosely connected manner.</p> <p><input type="checkbox"/> The text is about a highly familiar topic that is relevant to daily life.</p> <p><input type="checkbox"/> The text is not complex and has a predictable pattern of presentation.</p> <p><input type="checkbox"/> The texts is primarily written in present and present progressive tense.</p>	<p><input type="checkbox"/> The text provides a straight-forward sequencing of real-world events.</p> <p><input type="checkbox"/> The information in the text is conveyed through main idea and supporting details, and sentences are neatly connected through conjunctions and transitional expressions.</p> <p><input type="checkbox"/> The text is uncomplicated and the subject matter pertains to real-world topics.</p> <p><input type="checkbox"/> The comprehension of the text is supported by the reader's knowledge of the conventions of the language and Turkish prose.</p> <p><input type="checkbox"/> The text includes more complex grammatical structures such as passive voice, past and future tense.</p>

NEXT STEPS IN THE TURKISH C-TEST PROJECT



The Russian C-test Project

THE RUSSIAN C-TEST PROJECT

Goal

- To develop valid and reliable Russian language C-tests aligned with the American Council on Teaching of Foreign Languages (ACTFL) and Interagency Language Roundtable (ILR) proficiency guidelines

New Features

- Russian language C-test
- Subject-matter expert passage evaluations
- Think-aloud protocols for NSs and NNSs

RUSSIAN LANGUAGE FEATURES

Challenges

- Cyrillic alphabet
- Highly fusional morphology
- Verb conjugation for person, number, tense, voice, and mood
- Accent stress marker (') can distinguish between two different words



RESEARCH STAGES



CONCLUSIONS & FUTURE DIRECTIONS

- Continuing research
 - Additional languages
- Explore uses (remember ILTA Code of Ethics and Guidelines for Practice)
 - Research
 - Screening
- AELRC C-test Repository (in progress)
 - Bangla, French, Japanese, Portuguese, Spanish, Arabic, Turkish
 - To be made available to foreign language researchers and instructors affiliated with a university, educational or research institution
 - Must keep C-tests both available and relatively secure

THANK YOU! 谢谢 !

Teşekkür ederim! Спасибо!

Margaret Malone: malonem@georgetown.edu

Yiran Xu: yx110@georgetown.edu

Seyma Toker: st977@georgetown.edu

Derek Reagan: dr862@georgetown.edu



GEORGETOWN UNIVERSITY

