Measuring complexity and accuracy in L2 Korean oral production

AATK, June 27, 2015
Kyae-Sung Park
KOREAN, Department Of East Asian Languages & Literatures
University Of Hawai‘i
Contents

1. Purposes
2. Methods of the present study
3. Evaluations of complexity measures
   i. Basis analysis units
   ii. Morpho-syntactic complexity metrics
   iii. Lexical complexity metrics
4. Accuracy
5. Results
6. Limitations and implications
Groups for cross-study comparisons

- **English studies**
  - L2-English adults
  - L1-English adults
  - L2-English children
  - L1-English children

- **Korean studies**
  - L2-Korean adults
  - L1-Korean adults
  - L1-Korean children
Purposes

To provide a suitable measure of L2 proficiency for:

- **cross-age** comparisons for research comparing child L2 development with adult L2 development

- **cross-study** comparisons for research comparing lower-level L2ers with higher-level L2ers within the same population group
How to measure L2 proficiency (in research)

Thomas (1994): a survey of 157 articles from four acquisition journals

- L2 proficiency tends to be measured:

  a) Impressionistic judgment
     - lacking generalisability: one person’s ‘advanced’ is not another’s (Thomas 1994)
  b) Institutional status
  c) In-house assessment instrument
  d) Standardized test
     - often for one particular age group, e.g.
     - English: TOEFL for adults only; Korean: TOPIK for adults only
NOT suitable in the present study

- Adult L2ers attended several different schools (e.g. University of Hawaii, University of Washington, Flagship program at the University of Hawaii, etc.) and had followed different courses.

- No common denominator with which participants could be compared.
CA(F)-based L2 proficiency measures

*Complexity* and *accuracy*, together with fluency (CAF), have commonly been used to measure learners’ performance on tasks in L2 acquisition (e.g. Chaudron & Parker 1990; Larsen-Freeman & Long 1991; Robinson 2007; Skehan & Foster 1997; among others).
Defining L2 proficiency

“the ability to produce lexically, morphologically and syntactically complex and accurate utterances in the target language”

(Unsworth 2005: 153)
Figure 1. Composition of the construct of L2 proficiency in the present study
Present Study
Selection of a task

- Song & Schwartz (2009)
- Unsworth (2005, 2008)

- A picture-narration task was used to elicit the oral production data.
Procedure & Materials

- Participants were shown pictures on a laptop computer via PowerPoint and were asked to tell a story in their L2 about what they saw.

- Utterances were recorded using either a digital audio recorder (e.g. Zoom H2 portable stereo recorder) or the audio editing application Audacity.

- 5 to 10 minutes for each participant to complete the task
Data coding

- The recorded oral data were transcribed in a Word document format, and then exported into a txt file to be converted to CHAT transcription format using the CLAN (Computerized Language Analysis) program of CHILDES (MacWhinney, 2000).

- Each recording was transcribed 3 times by 3 transcribers in each language.

- The 3 transcriptions were cross-checked with each other, and when there was disagreement, I either made decisions based on the transcriptions and the oral data or checked it with the 2 transcribers to reach an agreement.
Participants

English study
- 30 L2-English adults
- 20 L1-English adults

Korean study
- 38 L2-Korean adults
- 9 L1-Korean adults
## Participants (L1ers)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Female : Male</th>
<th>Age at Time of Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1-English Adult</td>
<td>20</td>
<td>10:10</td>
<td>18–65 (M = 28.9)</td>
</tr>
<tr>
<td>L1-Korean Adult</td>
<td>9/(20)</td>
<td>18:2</td>
<td>21–41 (M = 28.2)</td>
</tr>
</tbody>
</table>

Table 1. L1 participants’ backgrounds in the present study
Participants (L2ers)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>F:M</th>
<th>Age at Time of Testing</th>
<th>Age of L2 Onset</th>
<th>Length of Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2-English Adult</td>
<td>30</td>
<td>20:10</td>
<td>18-50 (M = 28.2)</td>
<td>10–15 (M = 11.7)</td>
<td>0;2–22;0 (M = 3;8)</td>
</tr>
<tr>
<td>L2-Korean Adult</td>
<td>38/(39)</td>
<td>23:16</td>
<td>19-39 (M = 24.2)</td>
<td>15–33 (M = 20.0)</td>
<td>0;0–13;0 (M = 1;2)</td>
</tr>
</tbody>
</table>

Table 2. L2 participants’ backgrounds in the present study
Figure 1. Composition of the construct of L2 proficiency in the present study
Basis units of analysis
Basis units of analysis

a) Minimal terminal unit [T-unit]
b) Communication unit [c-unit]
c) Analysis of Speech unit [AS-unit]
T-unit

- “one main clause plus whatever subordinate clause and nonclausal expressions are attached to or embedded within it” (Hunt 1970: 14).

1. There are many different contributions between artists and scientists to society.  
2. First artists contribute to society for entertainment.  
3. Many people need it for relax after hard work.  
4. Artists contribute to society as film artists, singers and so on.  
5. Furthermore artists contribute to society with make new-work fields which are related to kind of activity.  

(Schneider & Connor 1990: 415)
T-unit (cont.)

- originally devised for L1 children’s writing

- popularly used for both spoken data and written texts in L2 research (e.g. Chaudron & Parker 1990; Larsen-Freeman 1978a, b, 1983; Larsen-Freeman & Strom 1977; Scott & Tucker 1974; Song & Schwartz 2009; Unsworth 2005).

- appropriate for written data produced by intermediate and advanced L2ers because they are likely to contain full clauses and sentences (Norris & Ortega 2009)
c-unit

- proposed in an attempt to capture the *elliptical nature of spoken data*

- grammatical independent predication(s) or ... answers to questions which lack only the repetition of the question elements to satisfy the criterion of independent predication... “Yes” can be admitted as a whole unit of communication when it is an answer to a question such as “Have you ever been sick?”  (Loban 1966: 5–6)

- “an independent grammatical predication, the same as a T-unit except that in oral language elliptical answers to questions also constitute predication”  (Chaudron 1988: 445)
In the question–answer pair, in which the answer *on the table* is not accompanied by a verb yet it has a communicative value, it is considered as a c-unit.

(2) Q: Where’s my hat?  
A: On the table.  
(Crooks 1990: 184)

(2’) Q: 내 모자 어디 있지?  
A: 책상에.
c-unit (cont.)

- c-unit includes both a grammatical predication within a speaker’s turn and an answer, which could be either grammatical or ungrammatical, to an interlocutor’s questions.

- more suitable for oral or dialogic data elicited in low-proficiency learners, which may involve low formality and contain many nonsyntactic segments.
AS-unit

- devised by Foster et al. (2000) as an improved alternative for oral discourse segmentation.

- “a single speaker’s utterance consisting of an independent clause or sub-clausal unit, together with any subordinate clause(s) associated with it” (Foster et al. 2000: 365).
AS-unit

(3)  
A:  
{which which} what is your opinion?  
1 AS
B:  
{maybe} {er} {he}
A:  
long time?  
1 AS
{or it’s for} for you it’s a major mistake or a small mistake?  
1 AS
B:  
maybe three months  
1 AS
A:  
three month for this one  
1 AS
okay  
1 AS
for me it’s ten  
1 AS
B:  
ten?  
1 AS
A:  
ten years  
1 AS
B:  
yeah  
1 AS
ten years  
1 AS
oh very long  
1 AS  
(Foster et al., 2000: 370, (48))
AS-unit (cont.)

- appropriate for highly interactional data that include many nonsyntactic segments
**T-unit** as basis unit of analysis

a) The data were elicited as a *monologue* without any interaction with or interruption by an interlocutor

→ **very few fragmentary and elliptical segments**

b) L2ers in this study fall into the intermediate and advanced levels in terms of morpho-syntactic and lexical accuracy

c) they produced full clauses and sentences most of the time

d) L2 spoken data are messier than written data, and application of T-units to segmentation requires specific guidelines.
Morpho-syntactic complexity measures
<table>
<thead>
<tr>
<th>Central Focus of Calculation</th>
<th>Measures</th>
<th>Key References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Length</strong>&lt;br&gt;(in morphemes, words, etc.)</td>
<td>• Mean length of utterance (MLU)</td>
<td>Brown (1973)</td>
</tr>
<tr>
<td></td>
<td>• Mean length of T-unit</td>
<td>Hunt (1965)</td>
</tr>
<tr>
<td></td>
<td>• Mean length of c-unit</td>
<td>Loban (1976)</td>
</tr>
<tr>
<td></td>
<td>• Mean length of clause</td>
<td>Scott (1988)</td>
</tr>
<tr>
<td><strong>Rate of Subordination</strong></td>
<td>• Mean number of (subordinate) clauses per T-unit</td>
<td>Elder &amp; Iwashita (2005)</td>
</tr>
<tr>
<td></td>
<td>• Mean number of (subordinate) clauses per c-unit</td>
<td>Skehan &amp; Foster (2005)</td>
</tr>
<tr>
<td></td>
<td>• Mean number of (subordinate) clauses per AS-unit</td>
<td>Michel et al. (2007)</td>
</tr>
<tr>
<td><strong>Verbal Density</strong></td>
<td>• Mean number of verbs per T-unit</td>
<td>Chaudron &amp; Parker (1990)</td>
</tr>
<tr>
<td></td>
<td>• Mean number of verbal utterances and modals, semi-auxiliaries, infinitival to, and tensed verbs per utterance</td>
<td>Valian (1991)</td>
</tr>
</tbody>
</table>

Table 3. Morpho-syntactic complexity measures in previous research
Length-based metrics

Length-based metrics = \frac{\text{the number of morphemes or words}}{\text{the number of chosen analysis units}}
Length-based metrics (cont.)

- MLU is widely used as a measure of morpho-syntactic development in L1 acquisition (Brown 1973)

- a valid indicator of syntactic complexity in the earliest stages of linguistic development up to a value of MLU 3.

- continues to develop after age five (Minifie et al. 1963; Loban 1976; Hunt 1970)

- highly variable in adult native speakers

→ unsuitable measure of grammatical complexity
Length-based metrics in L2 acquisition

Not appropriate in L2 acquisition research (Dewaele 2000)

- Subject to considerable variability
  
  a) Lack of systematicity: the most advanced L2ers did not produce the longest utterances (Lightbown, 1977)
  
  b) MLU is not linear (Blum-Kulka Olshtain: 1986)
     - In making requests, low proficiency L2ers use shorter utterances than native speakers, while high proficiency L2ers use longer utterances than native speakers
“L2ers are usually capable of producing multi-word/morpheme sentences almost immediately after initial contact with the target language” (Adamson 1988; Larsen-Freeman & Strom 1977:124).

“L2ers’ initial MLU is comparatively high and hence there is less room for the L2er to develop in this respect” (Unsworth 2008: 16)
Mean length of errorfree T-unit is the most satisfactory measure of grammatical complexity in L2 development (Larsen-Freeman 1983)

Less than ideal (Unsworth 2005)

- the number of errorfree T-units is so low that any assessment of proficiency based on it might not be truly representative of the participants’ L2 competence because substantial proportion of their total output would have to be discarded
Rate of subordination

\[
\text{Rate of subordination} = \frac{\text{the total number of (subordinate) clauses}}{\text{the number of chosen analysis units}}
\]

- Although subordination metrics have been widely used as a measure of complexity, they have received criticism as a measure of complexity in developing children, because it increases with increasing age in older L1 children, like MLU does (e.g. Hunt 1970; Loban 1976).
Rate of subordination (cont.)

- Unsworth (2005): study with L1 children aged 7–11 finds that “no fewer than ten of the 30 children produce NO subordinate clauses whatsoever” (p. 183).

- If one were to rely on subordination measures, the target native-like level for L2 participants would be very low, which consequently leads to very little room to measure development.
Rate of subordination (cont.)

- One third of the L1 children (6 out of 18) produced no subordinate clauses at all.
- The mean score was 0.11 for the L1 children and 0.28 for the L1 adults.
- Subordination is unsuitable for measuring syntactic complexity
Verbal density

Verbal density = \frac{\text{the number of (finite and nonfinite) verbs}}{\text{the number of chosen analysis units}}

- Originally used as a measure of syntactic complexity (Pica & Long 1986) to analyze ESL teacher speech as well as conversations between native speakers and nonnative speakers.

- Chaudron & Parker (1990) adopted it to measure proficiency using L2ers’ free and elicited-production data.
Verbal density (cont.)

- Valian (1991) measured modals, semi-auxiliaries, and tensed verbs, verbs (i.e. pure transitive, pure intransitive, optionally transitive) per utterance in 2 year-old L1-English.

- Children’s use of modals, tensed verbs, and (tensed and tenseless) verbs increases as MLU advances

- The usage of modals and verbs are strongly correlated with overt-subject use

→ indicates that verbal density can be used as an index of morpho-syntactic complexity in developing learners.
Verbal density as the morpho-syntactic complexity

a) can capture grammatical complexity at the clause level by including both finite and nonfinite verbs.

b) can be calculated regardless of targetlikeness because it counts both tensed and tenseless verb forms, thereby including learners’ non-targetlike inflectional forms.

(Unsworth 2005)
Lexical complexity measures
Lexical complexity measures

- “lexical complexity” = “lexical diversity” or “lexical richness”

- type-token ratio (TTR)
- Guiraud’s root TTR (Guiraud’s index)
- corrected TTR
TTR

\[
TTR = \frac{\text{the total number of types}}{\text{the total number of tokens}}
\]

- widely used in the assessment of lexical diversity; yet, it is well-known that TTR is susceptible to length of utterances.
- Richards (1987): the type-token ratio decreases as the number of tokens increases
- Johansson (2008): the reason for this is that a speaker/writer often has to use the same function words in order to produce one new (lexical) word, although the number of tokens can increase infinitely
Revised TTR

Guiraud’s Index = \frac{\text{the number of different lexical types}}{\sqrt{2 \times \text{the total number of tokens}}}

Corrected TTR = \frac{\text{the number of different lexical types}}{\sqrt{\text{the total number of tokens}}}

- By taking the square root of the total number of tokens, the TTR problem of the ratio being negatively correlated with sample size is alleviated; not independent of sample size, either.
- Richards (1987): the ratio is positively related to the number of tokens.
- Richards (1987): the ratio curve, which rises sharply as sample size increases, begins to level out when the number of tokens reaches approximately 250.
Accuracy
Accuracy

\[
\text{Accuracy} = \frac{\text{the number of error-free T-units}}{\text{the total number of T-units}}
\]

in terms of morpho-syntax and the lexicon
Criteria an error-free utterance in L2

a) When utterances include self-corrections, they are evaluated on the basis of the corrected form.

b) When the participant makes repeated errors, using a non-targetlike form of a verb stem or using a particular word incorrectly, only the first instance is counted as an error, because otherwise one persistent error can result in inappropriate deflation of the score.

c) As Korean does not have subject-verb agreement or an article system, there are no morpho-syntactic errors involving those syntactic properties in the L2-Korean data.

d) Korean is a null-argument language and therefore arguments can be dropped when they are recoverable from the discourse context. Moreover, case can be omitted especially in a colloquial language. Accordingly, omission of arguments or case markers is not considered as an error.
Inter-rater reliability

- Each transcription was checked for accuracy twice by 4 native-speaker student assistants in English and 2 native-speaker student assistants in Korean.

- The three judgments were cross-checked with each other.

- When there was a disagreement, in some cases I checked with the raters and in other cases, especially for Korean, I made my own decision based on the annotations provided by the raters.
Inter-rater reliability

- To assess inter-rater reliability, Cohen’s Kappa was conducted.

- A sample of 10% of the total transcripts was randomly selected.

- Cohen’s Kappa was found to have a value of 0.837 for the English data and 0.813 for the Korean data.

- Kappa has a range from 0 to 1.00; any Kappa value larger than 0.81 is considered to indicate almost perfect agreement (Landis & Koch, 1977).
Results: T-units
English data

<table>
<thead>
<tr>
<th></th>
<th>Coder A</th>
<th>Coder B</th>
<th>Coder C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coder A</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coder B</td>
<td>.947***</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Coder C</td>
<td>.922***</td>
<td>.977***</td>
<td>—</td>
</tr>
</tbody>
</table>

*** $p < .001$ level.

Table 5. Correlations of number of T-units among coders for the L2-English transcripts
Korean data

<table>
<thead>
<tr>
<th></th>
<th>Coder A</th>
<th>Coder B</th>
<th>Coder C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coder A</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coder B</td>
<td>.957***</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Coder C</td>
<td>.965***</td>
<td>.969***</td>
<td>—</td>
</tr>
</tbody>
</table>

*** $p < .001$ level.

Table 6. Correlations of number of T-units among coders for the L2-Korean transcripts
## Results of T-units

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>English Data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 Adults</td>
<td>20</td>
<td>36.6</td>
<td>12 – 68</td>
</tr>
<tr>
<td>L2 Adults</td>
<td>30</td>
<td>28.3</td>
<td>16 – 50</td>
</tr>
<tr>
<td><strong>Korean Data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1 Adults</td>
<td>9</td>
<td>23.1</td>
<td>16 – 43</td>
</tr>
<tr>
<td>L2 Adults</td>
<td>39</td>
<td>20.7</td>
<td>11 – 67</td>
</tr>
</tbody>
</table>

Table 7. Number of T-units by participants in the picture-narration task in the present study
T-unit coding guidelines for English data

Ortega, L., Iwashita, N., Rabie, & Norris, J. (in prep.)

1.1 Count **and**, **but**, **so**, and **though** as coordinate clauses (separate T-units).

— And they it come to it comes to them **and** it sits at the foot of her. (2 Ts)

— They have a children, **but** her mother and father the woman’s parents want live want to live with with them. (2 Ts)

— So he prefer only only thier love two love two person’s love, **but** she thinks family’s life. (2 Ts)

— woman has no money **so** she stole the baget from the bus and run away. (2 Ts)
T-unit coding guidelines for English data (cont.)

### 1.4 Count coordinate clauses where the second clause has a grammatical subject deletion as 1 T-unit. Unless the subject deletion reads/sounds clearly ungrammatical, count as 1 T-unit.

- the woman stand up and run away from the street. (1 T)
- So police arrested him and brought him go back brought him away. (1 T)
- The woman saw girl but didn’t see a man. (1 T)
- Two boys are firstly hestaty hestated, but soon become friendly with other children (1 T)
Korean is a null-argument language.

T-unit coding guidelines for Korean data

1.1 Count 고, 고서, (으)며, (으)면서, 지만, 거나 as coordinate clauses (separate T-units).

— 아 뭐 영희는 얼굴을 씻었고 Ø 뭐 아침 준비를 아 다 했습니다. (2 Ts)
— 그 애 영희는 옷을 입고 Ø 아 그 뭐 아 가방을 들고 아 학교에 가는 학교에 가기 전에 해야하는 준비를 합니다 네. (3 Ts)
— 그래서 어머니는 아 그 그 아들의 방으로 가서 Ø 아 그 재미있는 이야기를 읽었어요. (3 Ts)
— 그 그 젊은 곰은 아 괴찰을 것 같았지만 Ø 사실 다시 그 부모님 방으로 갔고 Ø 다른 일 때문에 일어났다고 그랬어요. (3 Ts)

- Classification between coordination and subordination based on Sohn (1999)
Results: Verbal density
Results of morpho-syntactic complexity (verbal density)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Range</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1-English Adults</td>
<td>20</td>
<td>2.167</td>
<td>1.571 – 2.976</td>
<td>0.404</td>
</tr>
<tr>
<td>L2-English Adults</td>
<td>30</td>
<td>2.160</td>
<td>1.500 – 2.906</td>
<td>0.379</td>
</tr>
<tr>
<td>L1-Korean Adults</td>
<td>9</td>
<td>1.653</td>
<td>1.313 – 2.045</td>
<td>0.242</td>
</tr>
<tr>
<td>L2-Korean Adults</td>
<td>38</td>
<td>1.440</td>
<td>1.056 – 1.875</td>
<td>0.203</td>
</tr>
</tbody>
</table>

Table 8. Results of verbal density in the present study
Crosslinguistic differences

so a little girl wearing green and green bows in her hair got up first thing in the morning and she decided to brush her teeth and next she decided to wash her face make it nice and clean and then she thought it would be a good idea to have a breakfast

- Verbal density: 11/4 = 2.75

어 어떤 여자애가 이제 학교 갈려고 아침에 이제 준비를 하고 있는데 아침에 이제 이빨도 닦고 거울 보면서 얼굴도 이제 깨끗하게 씻고 지금 그러고 있거든요. 그런 다음에 엄마가 이제 밥 먹고 가라 그래 가지고 아침에 이제 시리얼에 우유 타서 아침 맛있게 먹고 그 다음에 예쁘게 분홍색 원피스도 입고

- Verbal density: 11/8 = 1.38
Results: Guiraud’s index
### Results of lexical complexity (Guiraud’s Index)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Range</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1-English Adults</td>
<td>20</td>
<td>7.989</td>
<td>4.936 – 10.607</td>
<td>1.441</td>
</tr>
<tr>
<td>L2-English Adults</td>
<td>30</td>
<td>6.636</td>
<td>5.170 – 8.379</td>
<td>0.758</td>
</tr>
<tr>
<td>L1-Korean Adults</td>
<td>9</td>
<td>7.875</td>
<td>5.625 – 10.145</td>
<td>1.353</td>
</tr>
<tr>
<td>L2-Korean Adults</td>
<td>38</td>
<td>6.990</td>
<td>4.817 – 10.473</td>
<td>1.284</td>
</tr>
</tbody>
</table>

Table 9. Results of Guiraud’s index in the present study
Summary of results
Results of T-units and complexity measures

- **T-units**: Both L2 groups produced fewer T-units than the native-speaker groups did.

- **Verbal Density**: the mean score of the L2 Korean group was lower than that of the Korean native-speaker control group, while the L2 English group’s scores were similar to the English native-speaker control group’s. The mean scores and maximum scores of the L2-Korean learners were lower than those of the L2-English learners.

  → Due to crosslinguistic variation between Korean and English

- **Lexical complexity**: No such crosslinguistic variation was found for lexical complexity.
Results: Accuracy
## Results of Accuracy

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Range</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1-English Adults</td>
<td>20</td>
<td>95.2%</td>
<td>83.3% – 100%</td>
<td>0.051</td>
</tr>
<tr>
<td>L2-English Adults</td>
<td>30</td>
<td>68.2%</td>
<td>30.8% – 95.0%</td>
<td>0.173</td>
</tr>
<tr>
<td>L1-Korean Adults</td>
<td>9</td>
<td>99.2%</td>
<td>95.5% – 100%</td>
<td>0.016</td>
</tr>
<tr>
<td>L2-Korean Adults</td>
<td>38</td>
<td>61.1%</td>
<td>35.3% – 85.0%</td>
<td>0.148</td>
</tr>
</tbody>
</table>

Table 10. Results of rate of error-free utterances in the present study
Computation of L2 proficiency
Figure 2. The proficiency measurement overview in the present study
Table 11. Proficiency details by proficiency level in each group in the present study

<table>
<thead>
<tr>
<th>Group</th>
<th>Level</th>
<th>N</th>
<th>Mean</th>
<th>Range</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2-English Adults</td>
<td>Higher</td>
<td>16</td>
<td>1.211</td>
<td>-0.087 – 4.806</td>
<td>1.362</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>14</td>
<td>-1.384</td>
<td>-3.822 – -0.278</td>
<td>1.037</td>
</tr>
<tr>
<td>L2-Korean Adults</td>
<td>Higher</td>
<td>19</td>
<td>1.211</td>
<td>0.088 – 3.508</td>
<td>1.045</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>19</td>
<td>-1.211</td>
<td>-3.333 – -0.039</td>
<td>1.060</td>
</tr>
</tbody>
</table>
Limitations and Implications

- We need further work for more fine-grained coding guidelines for T-unit analysis in Korean and criteria for Korean verbal density with a careful consideration of morpho-syntactic features in Korean.

- The measurement of L2 proficiency with respect to morpho-syntactic and lexical complexity and accuracy allows us to make comparisons among different levels of L2 proficiency within the same population group, as well as between L2 adults and L2 children at approximately the same proficiency level.