

ASSESSING FLUENCY CONSTRUCT FROM A SECOND LANGUAGE ACQUISITION PERSPECTIVE: THE CASE OF TEEP SPEAKING TESTS

John Slaght

(University of Reading)



RESEARCH TEAM MEMBERS

- Parvaneh Tavakoli University of Reading
- Gill Kendon University of Reading
- Lisa-Marie Hunter Leeds University
- John Slaght University of Reading

SECOND LANGUAGE ACQUISITION & ORAL FLUENCY

THE CONSTRUCT OF L2 FLUENCY

Segalowitz (2010)

Cognitive fluency: efficiency of the operation of the cognitive mechanisms underlying performance

Utterance fluency:
Observable & measurable features of fluency

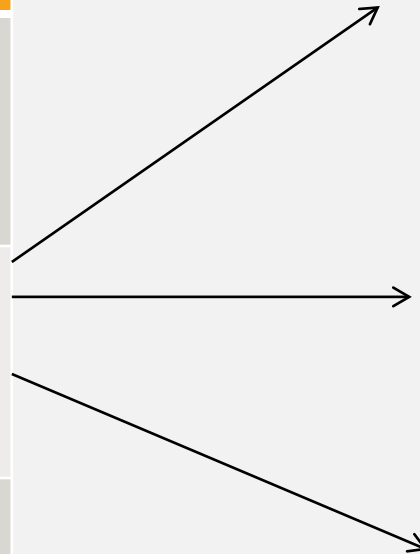
Perceived fluency: inferences listeners make about someone's cognitive fluency based on their perceptions of how fluent the speaker is

Skehan (2003)

Breakdown fluency (or how much silence is there)
Filled & silent pauses; mid or end-clause pauses; etc.

Speed fluency (or how fast speech is)
Speech rate; mean length of run, etc.

Repair fluency (or how many interruptions)
Repetition, hesitation, reformulation, false starts



**BACKGROUND
FLUENCY IN SPEAKING TESTS**

FLUENCY RATING SCALES

IELTS

SPEAKING: Band Descriptions

Band	Fluency and coherence	Lexical resource	Grammatical range and accuracy	Pronunciation
9	speaks fluently with only rare repetition or self-correction; any hesitation is content-related rather than grammatical; speaks coherently and develops full ideas	uses vocabulary with full range and precision; uses idiomatic language naturally	uses a wide range of grammatical forms flexibly and accurately	speaks with a clear, well-paced flow (fluid expression); speech is clear; may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility
8	speaks fluently with only occasional repetition or self-correction; any hesitation is content-related rather than grammatical; speaks coherently and develops full ideas	uses vocabulary with a wide range and precision; uses idiomatic language naturally	uses a wide range of grammatical forms flexibly and accurately	speaks with a clear, well-paced flow (fluid expression); speech is clear; may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility
7	speaks fluently with occasional repetition or self-correction; any hesitation is content-related rather than grammatical; speaks coherently and develops full ideas	uses vocabulary with a wide range and precision; uses idiomatic language naturally	uses a wide range of grammatical forms flexibly and accurately	speaks with a clear, well-paced flow (fluid expression); speech is clear; may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility
6	speaks fluently with occasional repetition or self-correction; any hesitation is content-related rather than grammatical; speaks coherently and develops full ideas	uses vocabulary with a wide range and precision; uses idiomatic language naturally	uses a wide range of grammatical forms flexibly and accurately	speaks with a clear, well-paced flow (fluid expression); speech is clear; may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility
5	usually speaks fluently with occasional repetition or self-correction; any hesitation is content-related rather than grammatical; speaks coherently and develops full ideas	uses vocabulary with a wide range and precision; uses idiomatic language naturally	uses a wide range of grammatical forms flexibly and accurately	speaks with a clear, well-paced flow (fluid expression); speech is clear; may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility
4	cannot respond with fluency; speaks with long pauses; has limited ability to link simple sentences; links basic sentences with simple connectives and some linking words	uses vocabulary with a limited range and precision; uses idiomatic language naturally	uses a limited range of grammatical forms flexibly and accurately	speaks with a clear, well-paced flow (fluid expression); speech is clear; may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility
3	speaks with long pauses; has limited ability to link simple sentences; links basic sentences with simple connectives and some linking words	uses vocabulary with a limited range and precision; uses idiomatic language naturally	uses a limited range of grammatical forms flexibly and accurately	speaks with a clear, well-paced flow (fluid expression); speech is clear; may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility
2	pauses frequently; speaks with long pauses; has limited ability to link simple sentences; links basic sentences with simple connectives and some linking words	uses vocabulary with a limited range and precision; uses idiomatic language naturally	uses a limited range of grammatical forms flexibly and accurately	speaks with a clear, well-paced flow (fluid expression); speech is clear; may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility
1	no communication; no rateable communication; does not extend beyond a few words	no communication; no rateable communication; does not extend beyond a few words	no communication; no rateable communication; does not extend beyond a few words	no communication; no rateable communication; does not extend beyond a few words
0	no communication; no rateable communication; does not extend beyond a few words	no communication; no rateable communication; does not extend beyond a few words	no communication; no rateable communication; does not extend beyond a few words	no communication; no rateable communication; does not extend beyond a few words

TRINITY COLLEGE LONDON
ISE II Speaking and listening rating scale

PTE Academic

TOEFL iBT® Test
Independent SPEAKING Rubrics

Score	GENERAL DESCRIPTION	DELIVERY	LANGUAGE USE
4	The response fulfills the demands of the task, with at most minor lapses in completeness. It is highly intelligible and exhibits sustained, coherent discourse. A response at this level is characterized by all of the following:	Generally well-paced flow (fluid expression). Speech is clear. It may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility.	The response demonstrates effective use of vocabulary and high-level grammar.
3	The response fulfills the demands of the task, with at most minor lapses in completeness. It is highly intelligible and exhibits sustained, coherent discourse. A response at this level is characterized by all of the following:	Generally well-paced flow (fluid expression). Speech is clear. It may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility.	The response demonstrates effective use of vocabulary and high-level grammar.
2	Speech is at an acceptable speed but may be uneven. There may be more than one hesitation, but most words are spoken in continuous phrases. There are few repetitions or false starts. There are no long pauses and speech does not sound staccato	Speech is at an acceptable speed but may be uneven. There may be more than one hesitation, but most words are spoken in continuous phrases. There are few repetitions or false starts. There are no long pauses and speech does not sound staccato	Speech is at an acceptable speed but may be uneven. There may be more than one hesitation, but most words are spoken in continuous phrases. There are few repetitions or false starts. There are no long pauses and speech does not sound staccato
1	Speech may be uneven or staccato. Speech (if >= 6 words) has at least one smooth three-word run, and no more than two or three hesitations, repetitions or false starts. There may be one long pause, but not two or more	Speech may be uneven or staccato. Speech (if >= 6 words) has at least one smooth three-word run, and no more than two or three hesitations, repetitions or false starts. There may be one long pause, but not two or more	Speech may be uneven or staccato. Speech (if >= 6 words) has at least one smooth three-word run, and no more than two or three hesitations, repetitions or false starts. There may be one long pause, but not two or more
0	Speech is slow and labored with little discernable phrase grouping, multiple hesitations, pauses, false starts, and/or major phonological simplifications. Most words are isolated, and there may be more than one long pause	Speech is slow and labored with little discernable phrase grouping, multiple hesitations, pauses, false starts, and/or major phonological simplifications. Most words are isolated, and there may be more than one long pause	Speech is slow and labored with little discernable phrase grouping, multiple hesitations, pauses, false starts, and/or major phonological simplifications. Most words are isolated, and there may be more than one long pause

Examiners often find the fluency criterion the most difficult to assess (e.g. Brown 2006b)
Research has shown that fluency is the most susceptible feature to elicitation tasks (e.g. Nakatsuhara 2012)

APPROACHES TO SPEAKING RATING SCALE DEVELOPMENT/VALIDATION

- **Empirical analysis of test-takers' speech samples** (e.g. Brown 2006a; Fulcher 1996; Fulcher, Davidson & Kemp 2011; Galaczi 2013; Nakatsuhara 2014; Turner & Upshur 1995)
- **Raters' perceptions of proficiency when rating spoken performances** (e.g. Brown 2006b; Brown & Ducasse 2009; May 2009; Orr 2002; Pollitt and Murray 1996)

De Jong's (2018) questions “the current conceptualization of fluency in language testing, in which it is defined as a concept that should be sought in the ear of the beholder, and where disfluency is only seen as a deficit.”

Fluency research: Language testing across assessed levels of proficiency

Fluency can predict proficiency

- Speed fluency (De Jong et al. 2012)
- Speed fluency and number of filled pauses (Revesz et al. 2014)
- Speech rate and mean length of run (Inoue 2013; Kahng 2014)

Tavakoli, Nakatsuhara & Hunter (2017):

- **RQ1:** How are various aspects of fluency presented across different **levels** of proficiency (A2, B1, B2, and C1) in the Aptis Speaking test?
- **RQ2:** To what extent is test-takers' fluency affected by **task** design?

Tavakoli, Nakatsuhara & Hunter (2017)

- (1) **Speed fluency distinguishes A2, B1 and B2 levels, but B2 and C1 levels are not different.**
- (2) **Length of silent pauses distinguishes A2 level from other levels.**
- (3) **Freq** **and B1)**
from
 - Sample size
 - Task issue (A2 level did not have performance on Task 4: an extended piece of speech)
 - Can the results be replicated?
- (4) **Freq** **B1 and**
C1
- (5) **Repair measures distinguish A2 and B1 levels; A2 produces very few and B1 most repairs**
- (6) **No effects of task type**

TEEP SPEAKING TEST

TEEP

- A **standardized/validated** English language proficiency test designed to examine academic abilities of students joining higher education in the UK (& around the world)
- Paired candidates + 2 examiners (interlocutor & assessor)
- It assesses test-takers from A1 to C2 (0-9); in practice the range is **B1 to C1 (4-8)**
- Overall time = 25 minutes; includes **planning time** before tasks

TEEP Speaking paper

Part	Task	Mode	Example	Planning time	Response Time
1	Individual Talk (role plays)	Monologue	Question: Which is better; private or public services	4 minutes	3 minutes
2	Scenario discussion	Dialogue	In pair, discuss with your partner and analyse the question	2 minutes	4 minutes
3	Focus question	Further discussion	Discuss the question further with your partner, and agree or disagree!	None	No time limit but generally about 2 mins

TEEP SPEAKING RATING SCALES

GLOBAL & ANALYTICAL CRITERIA

	Explaining ideas and information (Global criterion)	Interaction (Global criterion)	Fluency (Analytical criterion)	Accuracy, range (Analytical criterion)	Intelligibility (Analytical criterion)	
6.0	<i>Either detailed exploration of one or two given ideas OR covering all given ideas in a coherent, logical way OR using a relevant mix of given & own ideas.</i> Achieved by using some linking language and appropriate use of time available; BUT possible overuse of 'presentation' style formulaic phrases.	Engaged in interaction; able to compensate for lack of fluency. Turn-taking is usually suitable. Can initiate and build on own and partner's ideas intermittently, and respond to straightforward comments, etc. Can clarify or seek clarification, but may miss opportunities to do so. Some empathy.	Searches for words and hesitates at times but is reasonably fluent otherwise. May have occasional misunderstandings during oral communication. Any breakdown in communication will probably be self-corrected.	The candidate uses a range of structures and vocabulary, which allows the relatively easy exchange of ideas. Inaccuracies will occur fairly frequently but these errors do not usually obstruct communication. Few impeding errors, which may be self-corrected with some success.	Pronunciation idiosyncrasies related to prosodic and discrete-sound features may occasionally impede interaction. This may be evident through inability to pronounce certain words or sounds or through inappropriate intonation.	Competent speaker
5.0	Some or all of the given ideas communicated but with some difficulty; development is relevant but limited; efforts at explaining may be laboured or neglected; linking language may be repetitive and/or restricted; style not always appropriate	Basic interaction evident; topic superficially addressed; not always able to initiate or respond appropriately. Can follow and decode clearly expressed points. Some hesitations or lack of clarity can cause strain to listener. Lacks skills to clarify or ask for clarification. Only superficial consideration of partner's views.	Can communicate within a limited range of situations, either in a hesitant way OR by over-compensating for limitations by speaking too quickly or repetitiously. A breakdown in communication may occur.	The candidate can use a restricted range of structures and vocabulary fairly effectively. The exchange of simple ideas is usually problem-free, but the candidate will find it difficult to clarify more complex ideas. Errors will be frequent, a few may be impeding and self-correction may not be successful.	Pronunciation idiosyncrasies related to prosodic and discrete-sound features will probably impede interaction. The candidate may either not be able to pronounce a number of words or sounds and/or use unsuitable intonation.	Modest speaker

RESEARCH QUESTIONS

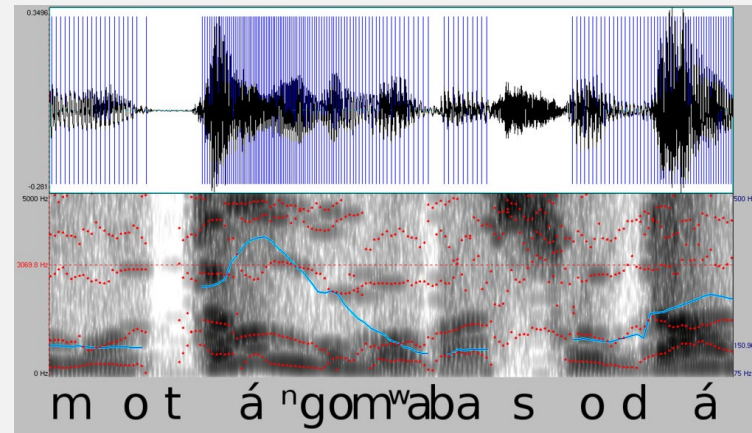
- The gap in the literature:
- Can the results of Tavakoli et al., (2017) be replicated with a larger sample and a different task?
 - Larger sample size
 - Same task for all participants
 - Different task conditions
- **RQ:** How are various aspects of fluency presented across different levels of proficiency (5.0, 5.5, 6.5, and 7.5) in the TEEP Speaking test?

METHODOLOGY

MATERIALS

- **60 test-takers in total:** 15 test-takers who were awarded overall scores of 5.0, 5.5, 6.5, and 7.5
- 15 test-takers x 4 proficiency levels x 3 minutes, totaling 168 minutes of recordings
- Recordings of the test-takers were selected on the basis of their **overall Speaking score** across all tasks

**SPEECH ANALYSIS
-PRAAT
BOERSMA & WEENIK, 2013**



FLUENCY MEASURES

- **Speed**
 - Articulation rate(pruned): mean number of syllables per minute divided by mean amount of phonation time (excluding pauses)
 - Speech rate (pruned): mean number of syllables per minute divided by total time (including pauses)
 - Mean length of run (pruned): the mean number of syllables between two pauses
 - Phonation time ratio: time taken to perform the task (excluding pauses)
- **Break down**
 - Mean length of silent pauses per 60 seconds at mid-clause and end-clause positions
 - Mean number of silent pauses per 60 seconds at mid-clause and end-clause positions
 - Mean number of filled pauses per 60 seconds
 - Mean length of filled pauses per 60 seconds
- **Repair**
 - Mean number of partial or complete repetitions (per 60 seconds)
 - Mean number of self-corrections (per 60 seconds)
 - Mean number of false starts and reformulations (per 60 seconds)
 - Total number of repair measures (per 60 seconds)

RESULTS

Speed fluency

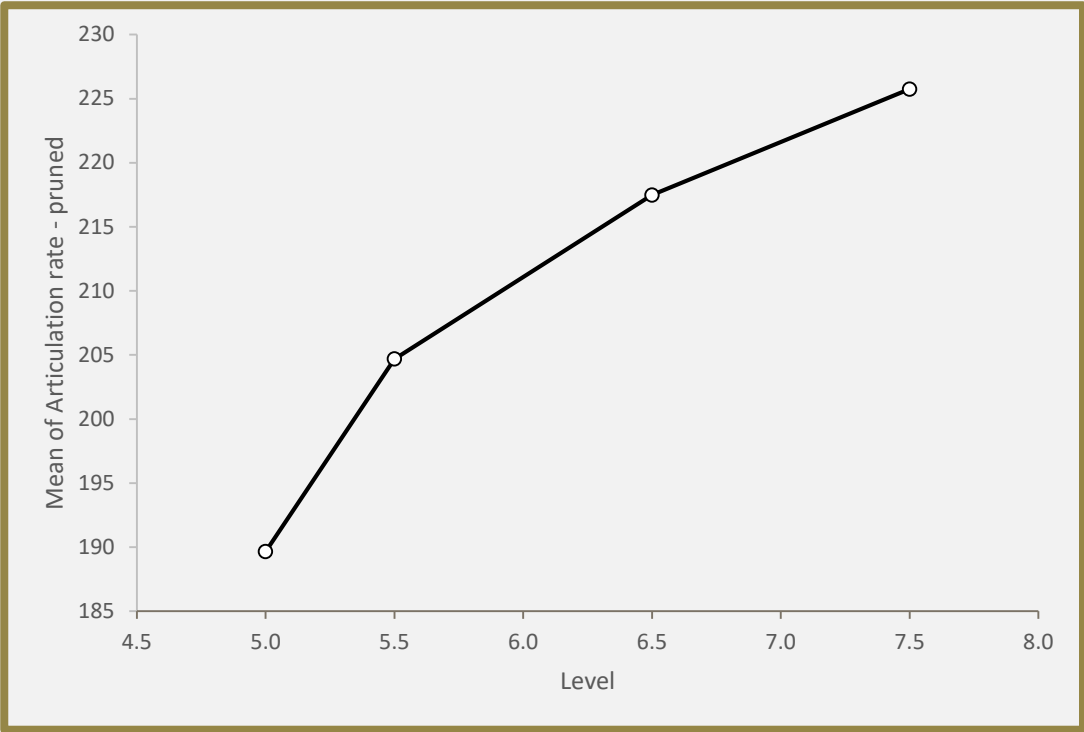


Figure 1: Articulation rate across proficiency levels



$(7.5=6.5) > 5.0; 7.5 > 5.5$

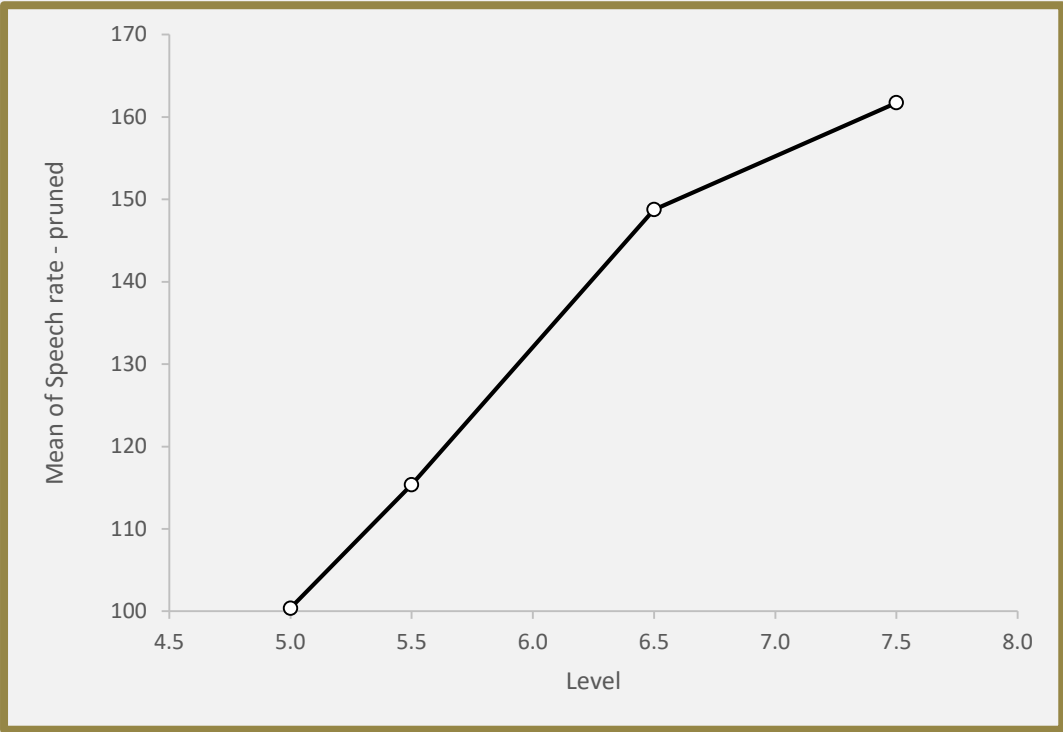


Figure 2: Speech rate across proficiency levels



$(7.5=6.5) > (5.5=5.0)$

Breakdown fluency: Frequency of pauses

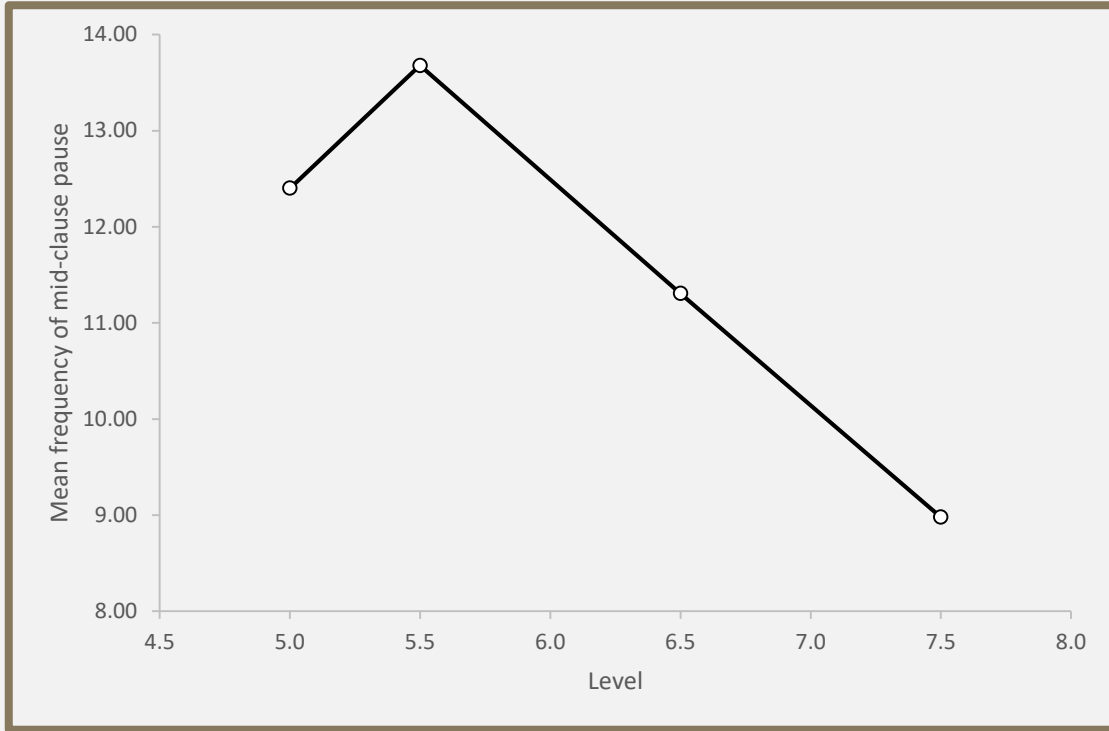


Figure 3: Frequency of mid-clause silent pauses

5.5 < 7.5

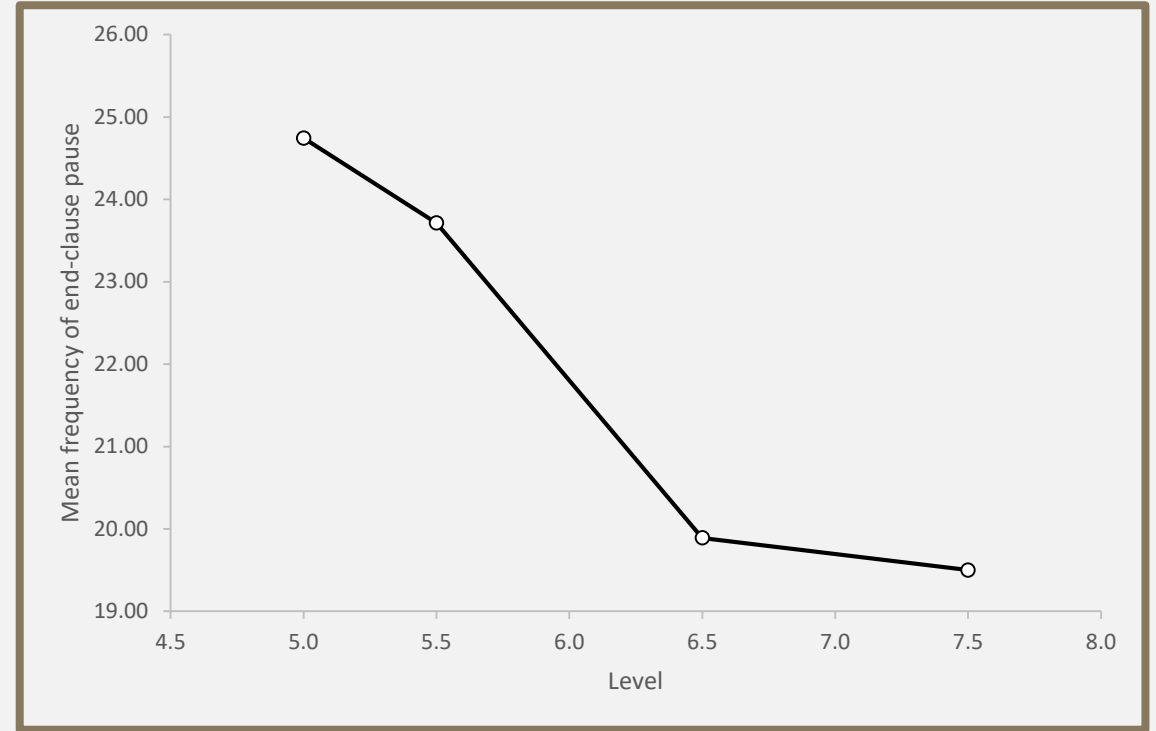


Figure 4: *Frequency of end-clause silent pauses*

**5.0 > 6.5, 7.5;
5.5 no difference from others**

Breakdown fluency: Length of end-clause pauses

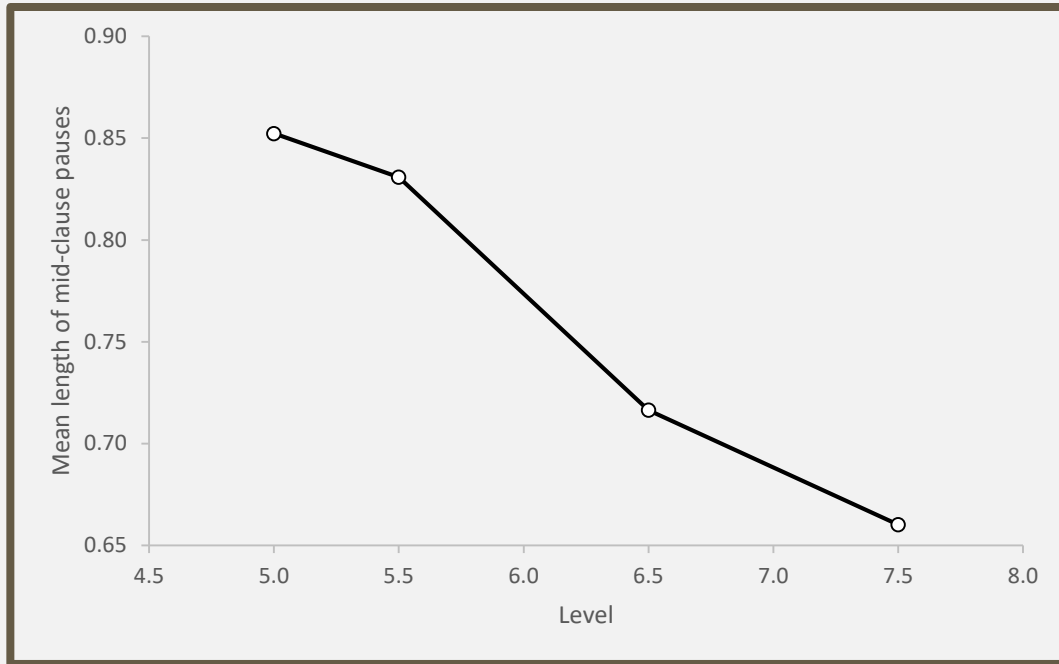


Figure 5: Length of mid-clause silent pauses

5.0=5.5 < 7.5; 6.5=7.5

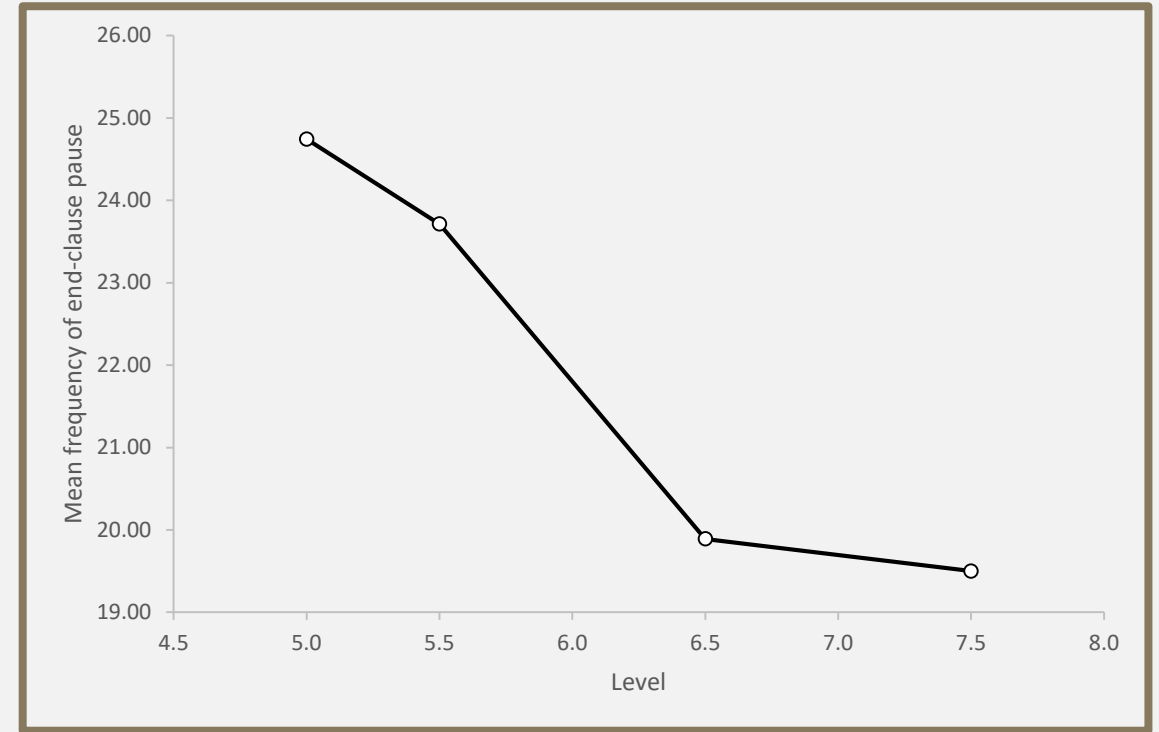


Figure 6: *Length of end-clause silent pauses*

**5.0 > 6.5, 7.5;
5.5 no difference from others**

Speed measures	Level
a) Articulation rate	(7.5=6.5) > 5.0; 7.5 > 5.5
b) Speech rate	(7.5=6.5) > (5.5=5.0)
c) Mean length of run	(7.5=6.5) > (5.5=5.0)
d) Phonation time ratio	(7.5=6.5) > (5.5=5.0)
Breakdown measures	Level
e) Frequency of mid-clause silent pauses	5.5 < 7.5
f) Mean length of mid-clause silent pauses	5.0=5.5 < 7.5; 6.5=7.5
g) Frequency of end-clause silent pauses	5.0 > 6.5, 7.5; 5.5 no difference from others
h) Mean length of end-clause silent pauses	5.0 > 6.5, 7.5; 5.5 no difference from others
i) Frequency of filled pauses	No statistically significant differences (5.5 produces filled pauses most frequently)
j) Mean length of filled pauses	No statistically significant differences (5.5 produced longest filled pauses)
Repair measures	No statistically significant difference

SUMMARY OF FINDINGS

- **Measures distinguishing proficiency levels**
- Speed fluency distinguishes 5.0 and 5.5 from 6.5 and 7.5 levels reasonably consistently. The two levels of 6.5 and 7.5 are not different in terms of speed fluency.
- Length of mid-clause silent pauses distinguishes 5.0 and 5.5 levels from 7.5 level. Length of end clause pauses distinguishes 5.0 from higher levels of 6.5 and 7.5.
- Frequency of mid-clause silent pauses only distinguishes 5.5 from 7.5 level. Frequency of end-clause silent pauses distinguishes 5.0 from 6.5 and 7.5.
- **Measures not distinguishing levels**
- Frequency of filled pauses
- Length of filled pauses
- Repair measures

CONCLUSIONS

Replicating Tavakoli et al. (2017)

Speed Fluency: distinguishing lower from higher levels

Speed Fluency: a ceiling effect

Length of silent pauses: distinguishing A2 from B2 and C1

Partial replication

Frequency of silent pauses: distinguishing A2 or B1 level from B2 or C1

Not replicated

Repair measures, filled pauses do not distinguish across level

- Repair measures affected by planning time?
- Filled pauses: a personal style?

THANK YOU!

