Annual Review of Applied Linguistics (2026), **46**, 144–178
doi:10.1017/S0267190526100269

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Sentence repetition, paraphrasing, and comprehension: Can you repeat a sentence without understanding it?

Scott Jarvis and Mary Akbary

Department of English, Northern Arizona University, Flagstaff, Arizona, USA
Corresponding author: Scott Jarvis; Email: scott.jarvis@nau.edu

Abstract

The present study addresses the question of whether an individual who does not understand a sentence might still be able to repeat it verbatim. To answer this question, we examined paraphrasing and repetition data from two previous studies: Pavlenko et al. (2019), which analyzed L1 and L2 participants' paraphrasing of seven Miranda warning sentences, and Akbary et al. (2023), which compared L1 and L2 participants' paraphrasing and elicited imitation (EI) performance on 30 commonly used EI sentences. We formulated our questions into four predictions, two of which directly addressed whether rote repetition without comprehension occurs at all. Our results confirmed both predictions, identifying 36 (5.6%) of the 646 instances of verbatim sentence repetition in the data as potential cases of repetition without comprehension. However, a broader analysis showed that the evidence for a lack of comprehension was relatively weak and ambiguous. We conclude with recommendations for overcoming the limitations of the present study and resolving the ambiguity of our findings.

Resumen

Esta investigación aborda la interrogante de si una persona que no entiende una oración podría aun así repetirla literalmente. Para responderla, examinamos datos de tareas de paráfrasis y repetición de dos investigaciones previas: Pavlenko et al. (2019), quienes analizaron cómo participantes de L1 y L2 parafraseaban siete oraciones de la advertencia Miranda, y Akbary et al. (2023), quienes compararon el desempeño de participantes de L1 y L2 en paráfrasis e imitación elicitada (EI) de 30 oraciones de uso común en tareas de EI. Nos planteamos cuatro predicciones, dos de las cuales abordaban directamente si la repetición mecánica sin comprensión realmente ocurría. Nuestros resultados confirmaron ambas predicciones e identificaron 36 instancias (5.6%) de las 646 repeticiones literales de oraciones en los datos como posibles casos de repetición sin comprensión. Sin embargo, un análisis más amplio mostró que la evidencia de falta de comprensión era débil y ambigua. Concluimos con recomendaciones para superar las limitaciones de la investigación actual y aclarar nuestros hallazgos.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Keywords: Miranda rights; Miranda warning; listening comprehension; elicited imitation; sentence paraphrasing

Introduction

A critical real-world language problem in the United States is that individuals who are arrested often do not understand their Miranda rights (e.g., Pavlenko et al., 2019). Forensic linguists (i.e., applied linguists serving as expert witnesses) are often hired by attorneys to provide expert analysis regarding whether the defendant is likely to have understood the Miranda warning before being questioned by police, as federal law requires (Miranda v. Arizona, 1966). Unfortunately, judges often dismiss linguistic experts' testimonies, ruling that their methods or conclusions do not meet the relevant evidentiary standards (see, e.g., McMenamin, 2022; Pavlenko, 2008). State and federal courts determine the admissibility of expert testimony primarily by applying standards such as the Frye test (*Frye v. United States*, 293 F. 1013 [D.C. Cir. 1923]), the Daubert standard (*Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 1993) – as interpreted through Federal Rule of Evidence 702 – or state-specific analogs to Rule 702. Each of these evidentiary frameworks – Frye, Daubert, and Rule 702 analogs – emphasizes the importance of methods that adhere to field standards and enjoy general acceptance in the relevant scientific community.

Unfortunately, no field standard yet exists for evaluating defendants' understanding of Miranda warnings, and there is ongoing disagreement among researchers and practitioners about whether comprehension is best assessed through paraphrasing tasks, repetition tasks, or other procedures (see, e.g., Akbary et al., 2023; Eggington & Cox, 2013; Pavlenko et al., 2019). At the same time, some forensic linguists have begun to treat a set of guidelines issued in 2015 by linguists, psychologists, and legal scholars (Communication of Rights Group [CoRG], 2015) as if it were a field standard (see, e.g., McMenamin, 2022). Importantly, these guidelines were developed for law enforcement officers delivering the Miranda warning rather than for forensic experts evaluating post hoc whether it was understood. The fact that the guidelines have been endorsed by the American Association for Applied Linguistics, the Linguistic Society of America, the International Association for Forensic and Legal Linguistics, and other scholarly associations, contributes to the perception that the guidelines represent a field standard whose recommendations are supported by past empirical research and are generally accepted by the field.

The present study relates to Recommendation 6 of the guidelines, which asserts that verbatim repetition of a sentence is not sufficient evidence of comprehension: "If suspects have difficulties restating the rights in their own words in English (e.g., *if they repeat the words just read to them* or if they remain silent), the interview should be terminated until a professional interpreter, with expertise in legal interpreting, is brought in" (emphasis added). The full recommendation maintains that paraphrasing (i.e., restating a sentence in one's own words) is the only acceptable test of Miranda comprehension. We recognize the intuitive logic of this assumption and are also aware of anecdotal evidence that individuals can repeat words and phrases in languages that

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

they do not know or understand. However, the guidelines do not include references to any research that supports the notion that an individual can repeat a complex sentence of the type that appears in Miranda warnings without understanding it. By contrast, several published, peer-reviewed studies have argued the opposite (e.g., Bley-Vroman & Chaudron, 1994; Wu et al., 2022) and have advocated for the use of sentence repetition as a measure of L2 Miranda comprehension in forensic contexts (e.g., Akbary et al., 2023; Benzaia et al., 2024; Eggington & Cox, 2013).

The simple yet multifaceted problem we address in the present study is whether verbatim repetition of a sentence is indeed possible in the absence of comprehension, as Recommendation 6 suggests. The multifaceted nature of the problem begins with the fact – reported to us by colleagues both privately and at academic conferences – that some forensic linguists cite Recommendation 6 of the CoRG guidelines as an authoritative field standard when justifying their use of sentence paraphrasing tasks as their method for assessing Miranda comprehension in court cases. If judges accept the notion that Recommendation 6 is a field standard and that paraphrasing is the only acceptable method for assessing L2 Miranda comprehension, then this could result in a situation where judges dismiss the methods and testimonies of forensic linguists who use any other method for assessing Miranda comprehension, such as elicited imitation (e.g., Benzaia et al., 2024; Eggington & Cox, 2013). If, however, the assumptions underlying Recommendation 6 are correct, then this needs to be confirmed, and this finding would inform the foundation of an eventual field standard for the forensic assessment of L2 Miranda comprehension. These are the problems and questions that the present study aims to address through a review of relevant prior research and a re–examination of the data from two recent studies that have used sentence paraphrasing tasks with relatively large groups of participants. The study aims to shed light on whether verbatim sentence repetition is possible when an individual has not understood the sentence.

Literature review

Sentence-level listening comprehension

Successful sentence-level listening comprehension requires the coordination of multiple knowledge components and processing skills (Goh & Vandergrift, 2021; Hogan et al., 2014). To understand a sentence, listeners need to be able to recognize (nearly) all the words in a sentence and be able to access their meanings rapidly (Bonk, 2000). Drawing on lexical, collocational, and grammatical knowledge, and supported by working memory, the listener incrementally combines word meanings and syntactic relationships into an integrated mental representation of the semantic proposition expressed by the stimulus sentence – a process that unfolds in real-time while the sentence is being heard (Horne et al., 2022). Failed sentence-level listening comprehension, conversely, results in a mental representation of the stimulus sentence that is either incomplete or inaccurate, or both (Goh, 2000; Meng et al., 2023).

Given that mental representations cannot be observed directly, the assessment of sentence-level listening comprehension relies on inferences made on the basis of listeners' responses to what they have heard. The most straightforward way to evaluate the content of the listener's working memory is to simply ask them to recall the sentence that they just heard (Daneman & Carpenter, 1980). Recalling the form of the sentence

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

entails sentence repetition, whereas recalling the meaning but not the exact form of the stimulus sentence involves either sentence paraphrasing or sentence translation.

Sentence paraphrasing as a measure of listening comprehension

The main advantage of sentence paraphrasing over sentence repetition is that a perfectly paraphrased sentence is unambiguous evidence of successful sentence comprehension: Could a person who did not understand the stimulus sentence possibly have produced such a complete and accurate paraphrase of it? (Acheson & MacDonald, 2009; Potter & Lombardi, 1990). Although a perfectly paraphrased sentence provides compelling evidence of comprehension, a poorly paraphrased sentence does not always provide compelling evidence of the lack of comprehension. The reasons for this are tied more to the nature of the task than to whether the listener's output constitutes a paraphrase or a repetition. Listeners who take the paraphrasing task seriously and who conscientiously endeavor to ensure that their output differs from the stimulus sentence have an increased cognitive load that goes beyond mere sentence comprehension and which may result in poor recall performance that underestimates their true ability to comprehend the sentence (Khrismawan & Widiati, 2013).

Similarly, the fact that paraphrasing requires not only receptive skills (i.e., decoding and comprehension) but also productive skills means that poor performance on a paraphrasing task might be the outcome of poor productive skills rather than comprehension failure, especially in cases where the person's productive skills lag behind their receptive skills. The effects of productive skills on recall performance have been demonstrated for elicited imitation tasks (e.g., Kim et al., 2016; Naiman, 1974) and are likely to be even greater for paraphrasing tasks given that the participant in a paraphrasing task is required to produce a response that differs from the stimulus sentence – thus requiring them to produce a novel linguistic formulation of the comprehended meaning while relying on their own productive knowledge of the language rather than simply repeating the words and structures given to them in the stimulus.

Perhaps the biggest disadvantage of sentence paraphrasing is the ambiguity that it creates for scoring. When individuals paraphrase, they often exercise discretion about what to include or omit, leaving out words they perceive as redundant, omitting details they judge as peripheral to the main point, removing parenthetical information or asides, and condensing what they perceive as repetitive phrasing (e.g., Akbary et al., 2023; Khrismawan & Widiati, 2013). Raters are sometimes unsure whether an incomplete paraphrase reflects a lack of comprehension or merely a deliberate strategy to make the paraphrased sentence more concise and focused than the original (e.g., Akbary et al., 2023; Wolfe & Goldman, 2005). An example of this provided by Akbary et al. (2023) is the following oral paraphrase, which was given by a participant in response to the stimulus sentence *Before he can go outside, he has to finish cleaning his room*: "He has to finish this room" (p. 11). The paraphrase omits information about what needs to be finished and what the consequences of doing so are, but it is not entirely clear whether the participant failed to understand these components of meaning or simply omitted what they considered to be implied by the context (e.g., "finish this room" in certain contexts means the same thing as "finish cleaning his room").

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

A recent systematic review of 157 peer-reviewed studies on L2 listening (Aryadoust & Luo, 2023) found no studies using paraphrasing as an assessment method, despite identifying multiple studies that used elicited imitation as a measure of general L2 listening proficiency. Our own search of the relevant literature identified only one empirical study outside the Miranda comprehension context: a small-scale study that used sentence paraphrasing to examine the effects of lexical and syntactic surprisal on listening difficulty (Mirzaei & Meshgi, 2020).

Within the context of Miranda comprehension, paraphrasing has been widely used by forensic practitioners – mainly psychologists – in cases where juveniles and those with mental disorders have been detained. Standard Miranda comprehension assessment instruments, such as the Miranda Rights Comprehension Instruments (MRCI), present detainees with sentences from the Miranda warning and ask them to paraphrase what they think they heard (Rogers & Drogin, 2019). Pavlenko et al. (2019) adopted the one-sentence-at-a-time paraphrasing approach from the MRCI while also following the CoRG guidelines' recommendation to have individuals "restate the rights in their own words." The study was designed as the first large-scale investigation of L2 speakers' ability to understand a common version of the Miranda warning, and it also aimed to determine whether Miranda comprehension can be expected to occur reliably beyond a certain level of overall L2 proficiency. The primary task in the study was a written Miranda paraphrasing task (i.e., listening to each sentence and then writing its meaning in their own words), which was administered to 59 L2 English participants and 41 L1 English participants. Due to concerns about cognitive load and task complexity associated with paraphrasing (Khrismawan & Widiati, 2013), the researchers also included a comparison group of 12 L2 English participants, who were given a Miranda dictation (i.e., written sentence repetition) task. The results showed that fewer than 3% of the L2 participants demonstrated a threshold level of Miranda comprehension, and that there was no level of proficiency at which Miranda comprehension problems were not found. Because the Miranda dictation results showed patterns similar to the paraphrasing results – including comparable score distributions under the same meaning-based rubric, regardless of whether responses were paraphrases or verbatim repetitions – the researchers concluded that the paraphrasing task did not underestimate participants' Miranda comprehension. Importantly, the researchers did not distinguish between successful paraphrases and successful verbatim repetitions, both of which were found in the participants' responses in both tasks. Distinguishing between the two types of responses is a central purpose of the present study.

To more fully examine the relationship between sentence paraphrasing and sentence repetition, Akbary et al. (2023) administered both an oral sentence paraphrasing and an oral sentence repetition (i.e., elicited imitation) task to 49 participants (L2 $n = 41$; L1 $n = 8$). Both tasks used the same set of 30 sentences taken from a widely used elicited imitation test (see Wu et al., 2022), which did not include Miranda warning sentences. The researchers examined whether the two tasks (involving different task instructions) elicited similar data and whether the two scoring rubrics produced parallel sets of scores. The answers to both questions were found to be affirmative, with the correlation between tasks (i.e., both scored with the same rubric) at $r = .96$, and the correlation between rubrics (i.e., with the same task scored with both rubrics) at $r = .95$.

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

The overall correlation between paraphrasing and elicited imitation (EI) when scored with their own rubrics was $r = .94$. Based on these results, the researchers concluded that the two tasks and the two rubrics measure the same construct: sentence-level listening comprehension.

We note, however, the possibility that the strong correlation might have also been due to an overlapping but not identical construct. Full construct validation would require not only an examination of convergent validity but also of divergent or discriminant validity (e.g., Campbell & Fiske, 1959) as well as a full cognitive explanation of the listening processes that L2 users engage in when performing EI, dictation, and paraphrasing tasks (e.g., Buck, 2001). As mentioned earlier, successful performance on both paraphrasing and EI tasks must also unavoidably involve productive skills. Both listening comprehension and language production furthermore engage a vast constellation of skills and components of knowledge related to phonology, lexis, grammar, discourse, pragmatics, and so forth (e.g., Buck, 2001). Thus, if paraphrasing and EI tasks are indeed measures of listening comprehension, then this is of course not all that they measure.

The claim by Akbary et al. (2023) that paraphrasing and EI tasks measure the same construct is nevertheless quite plausible given that the two types of tasks overwhelmingly elicit the same types of responses – that is, (successful or unsuccessful) paraphrases of the stimulus sentence – as we will describe in more detail in the empirical portion of the present study. Although less frequent, verbatim repetitions can also be found among participants' responses in both types of tasks. The primary differences between oral paraphrasing and EI are thus in the instructions given to participants – and what this implies about what participants are trying to accomplish – rather than in the general patterns of responses that they typically produce.

Beyond the questions of whether the two types of tasks measure the same construct and whether this construct is primarily listening comprehension, an additional conclusion of Akbary et al. (2023) was that the EI task provided more interpretable data than the paraphrasing task in the minority of cases where the two tasks did elicit qualitatively different data. The researchers found that participants were more likely to omit important meaning components when asked to paraphrase than when asked to repeat. They also found that the participants' EI responses more often provided unambiguous evidence of miscomprehension (e.g., hearing *light* as *line*). These differences appear to be attributable to differences in the task instructions, where paraphrasing instructions allow for a certain degree of creative license (e.g., "Restate the sentence in your own words") but where EI instructions do not (e.g., "Repeat the sentence exactly as you heard it"). Since forensic analysis in Miranda waiver cases centers on evidence of noncomprehension, the observed superiority of EI in eliciting such evidence led Akbary et al. to advocate its use in future work on the assessment of Miranda comprehension (see also Eggington & Cox, 2013).

Elicited imitation as a measure of listening comprehension

This recommendation was taken up by Benzaia et al. (2024), who investigated whether the findings of Pavlenko et al. (2019) could be confirmed by testing Miranda comprehension with an EI task (oral repetition) in place of the written paraphrasing task in

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Pavlenko et al. The 41 participants in the study by Benzaia et al. included 31 L2 speakers and 10 L1 speakers of English. The L1 speakers' score distribution in the two studies was nearly identical and led to approximately the same Miranda comprehension threshold for both (i.e., a score of 15 out of 22 in Pavlenko et al., and a score of 14 in Benzaia et al.). The study by Benzaia et al. confirmed that EI produces comparable results to sentence paraphrasing. This is true for both Miranda warning sentences (Benzaia et al., 2024) and the types of sentences typically included in EI tasks (Akbary et al., 2023).

Unlike sentence paraphrasing, EI has been used as a measure of language ability in numerous studies and has been the subject of a good deal of validation research. It is alternatively referred to as sentence repetition (Rujas et al., 2021), sentence recall (e.g., Ebert, 2014), and elicited oral response (Eggington & Cox, 2013). It is frequently used as a measure of children's language development (e.g., Ebert, 2014; Rujas et al., 2021) and L2 proficiency (e.g., Kostromitina & Plonsky, 2022; Park et al., 2020). Some studies have investigated which area(s) of L2 proficiency EI performance is most indicative of. Cox and Davies (2012) found that EI scores correlate more strongly ($r = .74$) with scores on the listening comprehension section of a standardized proficiency test than with scores on any other section. Wu et al. (2022) similarly found that EI scores correlate more strongly with Can-Do statements for listening ($r = .77$ and .80 for two different EI tests) than for other areas of language ability. Importantly, across the relevant studies, listening comprehension consistently emerges as one of the two strongest correlates of EI performance – both in terms of individual performance and group-level statistical relationships.

The other strong correlate, not surprisingly, is speaking ability. As we have discussed, successful performance on both EI and paraphrasing tasks requires not only strong listening skills but also well-developed language production skills (e.g., Kim et al., 2016). These two abilities go hand-in-hand in both EI and paraphrasing performance, such that a deficiency in either ability is likely to result in failed sentence recall. Some studies have found that listening ability is a stronger predictor of EI performance than speaking ability is (e.g., Cox & Davies, 2012; Wu et al., 2022), whereas other studies have found speaking ability to be the stronger of the two predictors (e.g., Bowden, 2016; Kim et al., 2016). Wu et al. (2022) "speculate[d] that listening abilities might be more important in EI performance for L2 users with more foreign-language-like learning experience, and speaking abilities may be more predictive of L2 users with more second-language-like, immersive learning experience" (p. 294). As an alternative and perhaps overlapping explanation, we believe that the relative strengths of listening vs. speaking skills as predictors of EI performance might depend on whether participants have speaking skills sufficient for expressing what they have understood. Further research will be needed to resolve this matter. In the meantime, the findings of the relevant research consistently point to EI scores as reliable indicators of both sentence-level listening comprehension and related speaking abilities.

Regardless of the group-level statistical patterns found in past research, which show that EI scores are strong general indicators of sentence-level listening comprehension, the fact that participants' EI responses are usually paraphrases rather than verbatim repetitions leaves open the question of the relationship between repetition and comprehension. For example, does a person's ability to repeat individual sentences always

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

reflect comprehension? Is it possible for a person to repeat a sentence verbatim without understanding it? A study by Frizelle et al. (2017) on 33 typically developing L1 English children concluded that this does happen. The study compared the children's performance on a sentence repetition task with their performance on a picture-based multiple-choice test using the same set of sentences. The researchers operationalized sentence comprehension as accurate performance on the multiple-choice items, and the results showed that the children were often able to repeat sentences that they did not comprehend. However, the researchers acknowledged in their interpretations that the comprehension test likely "involved skills over and above language knowledge" (p. 1452).

The clearest examples of repetition without comprehension come from the literature on nonword repetition. Studies on nonword repetition have found that L1 children's and adults' ability to repeat nonwords or sequences of nonwords improves with the word-likeness of the stimuli and with the participants' ability to chunk adjacent familiar-sounding syllables (e.g., Gathercole et al., 1991; Gupta & MacWhinney, 1997). However, even when these favorable conditions are present, successful nonword repetition becomes unlikely at five syllables and reaches its outer limit at approximately seven syllables, plus or minus two (e.g., Gathercole et al., 1991, 1994; Graf-Estes et al., 2007; Gupta & MacWhinney, 1997).

The findings of the literature on nonword repetition have been confirmed by the research on EI. Deygers (2020) found that L2 speakers can typically rely on rote repetition alone – without comprehension – to repeat short sentences up to eight syllables in length, but comprehension of the stimulus sentence is essential for correct repetition of sentences longer than approximately 10 syllables (see also Vinther, 2002). Studies by Kim et al. (2016), Okura and Lonsdale (2012), and Park et al. (2020) have similarly found that EI performance primarily reflects language proficiency and shows only weak or nonsignificant effects of memory capacity. To further mitigate the potential effects of phonological short-term memory, which is believed to decay within approximately 2 seconds after the stimulus (Baddeley et al., 1975), EI tasks are often designed with a pause of at least 2 seconds after the stimulus sentence is played until the participant is allowed to begin repeating it (Kostromitina & Plonsky, 2022). Overall, then, the existing literature on EI and nonword repetition unambiguously indicates that verbatim repetition is not expected and has not been found in statistical analyses of EI (or oral sentence repetition) data involving sentences longer than 10 syllables, especially when the participant is required to wait at least 2 seconds before responding. Nevertheless, the claims of Frizelle et al. (2017) combined with the overall lack of qualitative research in this area demonstrate that the question of whether verbatim repetition without comprehension might sometimes occur (in sentences exceeding 10 syllables) has not yet been satisfactorily answered.

The present study

Research questions and predictions

Our study is guided by two research questions that address the same general question: Does verbatim sentence repetition sometimes occur in the absence of comprehension? Our specific research questions are tailored to the available data and are intended to

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

represent the types of questions that might be raised by judges, attorneys, or opposing expert witnesses in Miranda waiver cases when determining whether a forensic expert's method meets the relevant evidentiary standards. Research Question 1 (RQ1) is most directly related to Recommendation 6 from the CoRG guidelines in that it addresses instances where a person has been asked to paraphrase a sentence but instead repeats it verbatim. We address this question through a re-examination of written paraphrasing data from Pavlenko et al. (2019) and oral paraphrasing data from Akbary et al. (2023). Research Question 2 (RQ2) is less directly related to Recommendation 6 but addresses the important question of the validity of EI data for testing sentence-level listening comprehension – specifically whether individuals might sometimes repeat a sentence verbatim on an EI task when their performance on a corresponding oral paraphrasing task shows that they are unlikely to have understood the sentence. The answer to this question will be valuable to forensic linguists who wish to use EI tasks to assess Miranda comprehension. We address RQ2 through a comparison of the oral paraphrasing and EI data from Akbary et al. (2023).

Our two research questions are operationalized into four predictions representing the four ways that the available data allow us to address the two research questions. Two of the four predictions (1a and 1b) are indirect, involving general tendencies that one would expect to find in the data if the assumptions underlying Recommendation 6 are correct. The more direct predictions (1c and 2) focus narrowly on whether individual participants may have repeated a sentence verbatim without having understood it fully. The predictions have been formulated to anticipate an affirmative answer to the research questions even though the literature generally suggests that the answer to both questions will be negative; our predictions are thus formulated to align with Recommendation 6 and to emphasize the types of evidence we are attempting to find.

Our research questions are as follows:

RQ1: When performing a paraphrasing task, do participants sometimes resort to verbatim repetition when the sentence is too difficult for them to comprehend?

> Prediction 1a (indirect): Effects of sentence difficulty on verbatim repetition. In the two sets of paraphrasing data, the occurrence of verbatim repetition will be higher for sentences that are more difficult for participants to comprehend.
> Prediction 1b (indirect): Effects of proficiency on verbatim repetition. In the two sets of paraphrasing data, the occurrence of verbatim repetition will be higher among participants who have more difficulty comprehending sentences.
> Prediction 1c (more direct): Individual patterns of verbatim repetition versus paraphrasing. In the two sets of paraphrasing data, the most difficult sentence that a given participant has repeated verbatim will be more difficult than the most difficult sentence that the same participant has unambiguously comprehended (i.e., paraphrased successfully).

RQ2: When performing both paraphrasing and EI tasks, do participants sometimes show an inability to paraphrase a sentence successfully on the paraphrasing task when they are able to repeat the same sentence verbatim on the EI task?

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Prediction 2 (direct): Individual patterns of verbatim repetition versus paraphrasing. In a comparison of participants' performance on the oral paraphrasing and EI tasks from Akbary et al. (2023), at least some participants will repeat a sentence perfectly on the EI task but will fail to paraphrase the same sentence successfully on the oral paraphrasing task.

Predictions 1a–1c rely on the notion of sentence comprehension difficulty and the assumption that participants might be able to produce exact repetitions of sentences that are too difficult for them to comprehend, as CoRG Recommendation 6 implies. The psycholinguistic explanation for these predictions is that a successful paraphrase reflects the ability to recognize and correctly interpret the meaning of each word, phrase, and grammatical construction in the stimulus sentence; the ability to incrementally combine these elements of meaning into an integrated mental representation of the semantic proposition expressed by the stimulus sentence; and the ability to hold this proposition in working memory while encoding it into a new sentence made up of words and/or grammatical constructions that differ at least partially from those of the stimulus sentence but which accurately express the original semantic proposition (e.g., Goh, 2000; Horne et al., 2022; Meng et al., 2023). Successful sentence repetition, on the other hand, is assumed for present purposes to reflect only phonological short-term memory (e.g., Baddeley et al., 1975), which does not require the participant to recognize, decode, or semantically integrate any of the words or grammatical constructions from the stimulus sentence.

Although the literature indicates that successful rote repetition is constrained by time (degrading within 2 seconds after the stimulus, e.g., Baddeley et al., 1975), length (limited to seven syllables, plus or minus two, e.g., Graf-Estes et al., 2007), and the degree to which the syllables in the stimulus sentence adhere to familiar phonotactic patterns (e.g., Gupta & MacWhinney, 1997), the greater demands of sentence paraphrasing versus rote repetition suggest that participants might still be able to repeat a sentence from rote memory even when the sentence is too difficult for them to paraphrase. Logically, the group-level consequences of this possibility are that (a) participants will resort to rote repetition more often with sentences that are more difficult to paraphrase (Prediction 1a), and (b) lower-performing participants who are less able to comprehend and paraphrase the stimulus sentences will show a greater tendency than higher-performing participants to rely on rote repetition (Prediction 1b). The sentence difficulty analysis related to Predictions 1a and 1b will also enable the identification of possible cases of repetition without comprehension – that is, instances where a participant has produced an exact repetition of a sentence whose difficulty exceeds the difficulty of all the sentences that the same participant has successfully paraphrased (Prediction 1c).

Prediction 2 takes advantage of the fact that the data from Akbary et al. (2023) include responses from a single set of participants who performed both oral paraphrasing and EI tasks with the same set of stimulus sentences. Even though Recommendation 6 applies to cases where a person has been asked to paraphrase a sentence but instead produces an exact repetition, whereas the relevant EI data from Akbary et al. (2023) include the responses of participants who were asked to repeat sentences verbatim and did as they were instructed, the data from Akbary et al. allow

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

us to examine whether at least some participants might have been able to repeat a sentence verbatim on the EI task while demonstrating an inability to paraphrase the same sentence successfully on the oral paraphrasing task. As with Prediction 1c, Prediction 2 entails the identification of any instance where a participant might have repeated a sentence verbatim (i.e., relying on rote repetition) without having understood the sentence. The relevant literature suggests that this will happen only with stimulus sentences that do not exceed 10 syllables in length (e.g., Deygers, 2020; Vinther, 2002), whereas Recommendation 6 does not explicitly assume such a constraint. In the present study, we will attempt to identify all instances where repetition without comprehension may have occurred and will further examine these instances in relation to sentence length, sentence difficulty, and the strength of the evidence that the participant failed to comprehend the sentence.

Method

Participants and data

We address the two RQs and four predictions through a careful analysis of three datasets taken from two previous studies: Pavlenko et al. (2019) and Akbary et al. (2023) (see Table 1). The dataset from Pavlenko et al. includes written paraphrases of seven Miranda warning sentences, whereas the two datasets from Akbary et al. consist of oral paraphrases and oral repetitions of a set of 30 ordinary sentences adopted from a commonly used EI test (see Wu et al., 2022). The data from the paraphrasing and EI tasks in Akbary et al. were produced by a single group of participants on separate days. In total, the data we analyze involve 150 unique participants, 37 unique stimulus sentences, and 3,700 unique sentence-level responses (2,200 responses from paraphrasing tasks, and 1,500 from an EI task). The first column in Table 1 provides a two-part identifier for each dataset, where the first part indicates the source it was taken from, and the second part indicates the task type (WP = written paraphrase, OP = oral paraphrase, EI = elicited imitation). The second column in Table 1 indicates the RQs and predictions that each dataset is used to address. The third column provides summary information about the task, stimuli, and participants represented in each dataset, and the final column cites the source of the dataset.

The participants in both studies included L1 and L2 speakers of English. The L2 speakers in Pavlenko et al. (2019) were international students in English language programs at various universities in the United States and the L2 speakers in Akbary et al. (2023) were immigrants to the United States, most of whom were enrolled in a free community language support program. The English proficiency levels of the L2 participants in the two studies ranged from approximately intermediate low to superior, with a somewhat higher overall distribution of proficiency levels for the immigrants than for the international students. The international students in Pavlenko et al. were L1 Arabic and L1 Chinese speakers, whereas the immigrants in Akary et al. were primarily L1 Spanish speakers ($n = 30$) but also included participants from nine additional L1 backgrounds. The L1 participants in both studies were adults with university educations. We note that Akbary et al. removed an L2 participant from their study due to the fact that he was found to be a statistical outlier in terms of his unusually low EI performance. However, we have included this participant in the present study so as not

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Table 1. Datasets analyzed in the present study

Dataset	RQs (Predictions)	Description	Source
Pavlenko_WP	RQ1 (1a–1c)	Task: Written paraphrase Stimuli: 7 Miranda sentences Participants: 100 (L1 = 41; L2 = 59)	Pavlenko et al. (2019)
Akbary_OP	RQ1 (1a–1c) RQ2 (2)	Task: Oral paraphrase Stimuli: 30 typical sentences Participants: 50 (L1 = 8; L2 = 42)	Akbary et al. (2023)
Akbary_EI	RQ2 (2)	Task: Elicited imitation Stimuli: 30 typical sentences Participants: 50 (L1 = 8; L2 = 42)	Akbary et al. (2023)

to exclude any available data that could potentially provide evidence of an instance of verbatim repetition without comprehension.

The written paraphrasing task in Pavlenko et al. (2019) required participants to listen to audio-recorded Miranda warning sentences, one at a time, and immediately after hearing each sentence (played only once), to restate the meaning of the sentence in their own words in writing on a piece of paper provided to them. Participants' responses were entered into a spreadsheet, separated into clauses, and scored on a scale of 0–2, where a score of 2 represented a fully adequate paraphrase with all major meaning components included. Scores of 0 and 1 represented unsuccessful attempts to paraphrase or repeat the clause. Importantly, participants sometimes repeated sentences on the paraphrasing task, and they often paraphrased sentences on the dictation task. The scoring protocol did not distinguish between successful paraphrases and successful repetitions – i.e., sentences repeated verbatim on the paraphrasing task were treated as successful paraphrases. Akbary et al. used a similar protocol for their oral paraphrasing task, except that they used 30 typical English sentences (originally from Ortega et al., 2002) instead of Miranda sentences, and each sentence was followed by a 2-second pause followed by a tone. Only after the tone were participants allowed to paraphrase the sentence, and they were asked to do so orally rather than through writing. Their responses were transcribed, separated into clauses, and scored with the same scoring protocol used in Pavlenko et al.

The EI task in Akbary et al. was administered to the same participants either 1–2 days before or 1–2 days after they were given the paraphrasing task. The EI task used the same 30 sentences in the same order as those used in the paraphrasing task. As in the paraphrasing task, sentences in the EI task were followed by a 2-second pause and an audible tone. After hearing the tone, participants were instructed to orally repeat the sentence exactly as they heard it. Responses were transcribed and scored according to the 5-point EI rubric proposed by Ortega et al. (2002), which has been used in multiple other studies (see, e.g., Wu et al., 2022). Unlike the paraphrasing rubric, which is applied to clauses, the EI rubric is applied to full sentences. The highest possible score for a response on the EI task is 4, which reflects exact verbatim repetition. A score of

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

3 represents a full and successful paraphrase of the stimulus sentence. Scores of 0–2 reflect unsuccessful paraphrases of the stimulus sentence. Importantly, most responses on the EI task were not verbatim repetitions, but scores of 3 allow us to identify which responses nevertheless reflect full comprehension.

We acknowledge substantial differences between the datasets from Pavlenko et al. (2019) and Akbary et al. (2023) in terms of the L2 populations they represent (international students vs. immigrants), task modalities (written vs. oral responses), stimuli (Miranda sentences vs. typical sentences), and cognitive demands (immediate restatement vs. 2-second delayed response). These differences preclude direct quantitative comparisons or statistical equating or anchoring of the data from the two studies. However, this is by design given the nature of our general research question, which is whether verbatim sentence repetition without comprehension can be found at all (i.e., in any population under any circumstances). Our inclusion of the data from these two studies is not for purposes of comparison but a result of the fact that these are the only relatively large-scale studies we are aware of that have tested L2 participants' sentence-level listening comprehension using a paraphrasing task. In accordance with Predictions 1a–1c, the participants in Pavlenko et al. might be expected to produce a greater number of rote repetitions given their lower distribution of L2 proficiency, the greater complexity of their stimulus sentences, and the fact that they were able to begin paraphrasing immediately after hearing each sentence rather than waiting until after the 2-second delay. Alternatively, the greater length of the Miranda sentences in Pavlenko et al. (*M*: 19.86 syll, range: 9–31 syll) might make rote repetition less likely in this dataset in comparison to that of Akbary et al. (*M*: 13.77 syll, range: 7–19 syll). Regardless of these differences, evidence of rote repetition without comprehension in the data from either study will support the predictions of the present study.

It is critical to recognize that repetitions and paraphrases can both be found among participants' responses on the paraphrasing and EI tasks. The primary difference between the two tasks is not in the responses but in the instructions given to participants (i.e., *restate the sentence in your own words* vs. *repeat the sentence exactly as you heard it*) and the implications that these instructions have for scoring. As we discussed earlier, paraphrasing instructions allow the participant a greater degree of discretion concerning which elements of meaning from the stimulus sentence to include in their response, whether to express these meanings explicitly (i.e., "finish cleaning the room" vs. "finish the room"), and whether to express them at a different level of semantic specificity (e.g., "has a wonderful sense of humor" vs. "is very funny") (Akbary et al., 2023). EI instructions, on the other hand, direct the participant to include all elements of meaning from the stimulus sentence. To the extent that they attempt to follow the task instructions, participants in a paraphrasing task focus on making their response different from the stimulus sentence, whereas participants in an EI task focus on making their response the same or as similar as possible to the stimulus. As a different-from-stimulus task, paraphrasing is logically more cognitively demanding than EI given that paraphrasing involves not just recall but also reformulation. Interestingly, however, data from the relevant literature (e.g., Akbary et al., 2023) show that repetition failures are more frequent than paraphrasing failures. We suspect that this is partly due to differences in the scoring process, where raters in a paraphrasing task might be less strict in their ratings than raters in an EI task given that paraphrasing both allows and

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

tacitly encourages creative responses. Despite these differences, the scoring protocols for both types of tasks treat verbatim repetitions and complete paraphrases as successful responses. In the present study, however, we will treat verbatim repetitions in both types of tasks as potential cases of rote repetition void of comprehension and will count only successful paraphrases as indicators of sentence comprehension.

Analysis

Prediction 1a. The analysis for this prediction focuses on the two paraphrasing datasets (i.e., Pavlenko_WP and Akbary_OP) and predicts that when asked to paraphrase a sentence, participants will resort to rote repetition more frequently with stimulus sentences that are more difficult to comprehend than with sentences that are easier to comprehend. Our predictor variable for this analysis is sentence comprehension difficulty, and the dependent variable is the occurrence of verbatim repetitions. Following the widely used EI scoring rubric proposed by Ortega et al. (2002), verbatim repetitions were identified as instances where the participant's response was identical to the stimulus sentence in that it consisted of the same words in the same order and with the same grammatical inflections. Punctuation and spelling were not taken into consideration. Verbatim repetitions that met these criteria were treated as potential rote repetitions given that rote repetition without comprehension relies on phonological short-term memory and results in the repetition of strings of sounds rather than strings of morphemes (e.g., Baddeley et al., 1975). Cases where the correct words were repeated without the correct inflectional morphemes reflect word recognition and morpheme-level processing rather than a strict phonological replay (e.g., West, 2012). Such cases were therefore treated as (successful or unsuccessful) paraphrases rather than as rote repetitions.

Sentence comprehension difficulty – our predictor variable for Prediction 1a – was operationalized solely in relation to paraphrasing performance after verbatim repetitions were separated out from responses that constituted (successful or unsuccessful) paraphrases. The separation of repetitions from paraphrases was done to avoid including the dependent variable (i.e., verbatim repetitions) in the predictor variable (sentence comprehension difficulty) and primarily because verbatim repetitions were treated as potential cases of rote repetition occurring in the absence of comprehension. Sentence comprehension difficulty was operationalized for each stimulus as the proportion of successful paraphrases among all paraphrase attempts for that stimulus (excluding verbatim repetitions). To identify successful paraphrases, we relied on the item scores from the original studies. However, because the paraphrasing tasks in Pavlenko et al. (2019) and Akbary et al. (2023) were scored at the level of individual clauses, we combined the clause-level scores for each sentence and identified as a successful sentence paraphrase any case where a participant received full points for each clause of a given sentence and where the participant's response was not a verbatim repetition of the stimulus sentence.

Although Prediction 1a posits a relationship between sentence difficulty (predictor) and verbatim repetition frequency (outcome), we forgo inferential tests: Past research indicates that rote repetitions without comprehension are rare and do not follow reliable statistical patterns. We therefore confirm Prediction 1a descriptively if the

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

sentence with the most verbatim repetitions ranks in the top third of stimulus sentences ranked by difficulty in the respective dataset (i.e., top 2 of 7 in Pavlenko_WP; top 10 of 30 in Akbary_OP). Although a confirmation of Prediction 1a in either dataset would provide compelling evidence of rote repetition without comprehension as a recurrent phenomenon, the relatively indirect way in which this prediction addresses RQ1 means that a rejection of Prediction 1a would not preclude the possibility that isolated cases of repetition without comprehension might still be found.

Prediction 1b. This prediction similarly focuses on the Pavlenko_WP and Akbary_OP datasets and predicts that the participants with the lowest levels of demonstrated comprehension – i.e., producing the fewest successful paraphrases – will show the highest reliance on verbatim repetition. The predictor variable, which we refer to for convenience as proficiency, was operationalized for each participant as the proportion of successful paraphrases among all paraphrase attempts by that participant (excluding verbatim repetitions). The dependent variable was operationalized as the number of verbatim repetitions produced by each participant across all stimulus sentences. Similar to Prediction 1a, we confirm Prediction 1b descriptively and liberally if the participant with the most verbatim repetitions ranks in the bottom third of participants ranked by proficiency (i.e., bottom 33 of 100 participants in Pavlenko_WP; bottom 17 of 50 participants in Akbary_OP). As before, given the relatively indirect way in which Prediction 1b addresses RQ1, a confirmation of Prediction 1b would support the possibility of rote repetition without comprehension, but a rejection of Prediction 1b would not rule it out.

Prediction 1c. This prediction also focuses on the two paraphrasing task datasets: Pavlenko_WP and Akbary_OP. In comparison to Predictions 1a–1b, Prediction 1c more directly addresses the question of whether any instances can be found of rote repetition without comprehension. The analysis for Prediction 1c relies on the available task-internal evidence of sentence comprehension and noncomprehension, which necessitates the use of contextual information from the two previous analyses regarding the difficulty levels of individual stimulus sentences and the patterns of comprehension (i.e., successful paraphrasing) found in the responses of individual participants. For the purposes of the present analysis, instances of rote repetition without comprehension are identified as cases where a participant has produced a verbatim repetition of a stimulus sentence that is more difficult to comprehend than any of the sentences that the same participant has paraphrased successfully. Sentence difficulty for this analysis is operationalized in the same way as for Prediction 1a. We confirm Prediction 1c with any instance where the most difficult sentence that a participant has repeated verbatim exceeds the difficulty of the most difficult sentence the same participant has paraphrased successfully, where the difficulty ranks of the two sentences are not adjacent (e.g., rank 25 for the successful paraphrase and rank 26 for the verbatim repetition). A straightforward confirmation of this prediction would provide strong support for the notion that rote repetition without comprehension is possible.

Prediction 2. The analysis for this prediction compares individual participants' responses in the paraphrasing and EI datasets from Akbary et al. (2023): Akbary_OP and Akbary_EI. The prediction is that at least some participants will repeat a sentence perfectly on the EI task but will fail to paraphrase the same sentence successfully on

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

the oral paraphrasing task. Verbatim repetitions, as before, are operationalized as exact repetitions, whereas successful paraphrases are operationalized as attempts at sentence paraphrasing where the raters from the original study assigned a score of 2 (i.e., full comprehension) to each clause of the participant's response. We confirm Prediction 2 as any instance where a participant has produced a verbatim repetition of the stimulus sentence in the EI task but has failed to produce a successful paraphrase of the same stimulus sentence in the paraphrasing task. Although a verbatim repetition in an EI task is somewhat different from the context described in Recommendation 6 (i.e., being asked to paraphrase a sentence but repeating it instead), a straightforward confirmation of Prediction 2 would constitute strong evidence that rote repetition without comprehension is possible.

To the extent that instances of this type are found in the data, we will examine the potential effects of sentence length, sentence difficulty, and any other considerations that might affect how the evidence should be interpreted. This is in alignment with several U.S. Supreme Court decisions that instruct judges to consider "the totality of the circumstances" when examining the evidence in Miranda waiver cases (e.g., *Davis v. United States*, 512 U.S, 1994; *Fare v. Michael C.*, 442 U.S, 1979). Of central relevance are the responses that reflect unsuccessful paraphrases. Some types of unsuccessful paraphrases provide stronger evidence of noncomprehension than others. The strongest evidence of failed comprehension includes minimal responses (e.g., "the person"), responses where the participant explicitly acknowledges a comprehension failure (e.g., "I don't know"), and responses that contradict the semantic content of the stimulus sentence (e.g., "The little boy who died yesterday is sad" as a response to the stimulus sentence *The little boy whose kitten died yesterday is sad*). Weaker, more ambiguous cases of failed comprehension include unsuccessful paraphrases that are either incomplete or at a different level of semantic specificity but semantically aligned with the stimulus sentence (e.g., "He says he wants a very good house which their dogs can live" as a response to the stimulus sentence "I want a nice, big house in which my animals can live"). While addressing the totality of the circumstances, we will assess the strength of the evidence for noncomprehension in any cases that meet our criteria for confirming Prediction 2.

Results

The results of the analysis for Prediction 1a are summarized in Tables 2 and 3. Table 2 shows the results of this analysis for the Pavlenko_WP dataset, where the seven Miranda warning sentences are ordered from most difficult to least difficult in accordance with the column labeled "New diffic. rank." The sentence IDs in the first column indicate the order in which the sentences were administered to the participants, and the second column shows the exact wording of each stimulus sentence. The third column shows the number of participants (out of 100) who repeated the sentence verbatim, and the fourth column shows the number of participants who paraphrased the sentence successfully (and did not repeat it verbatim). The fifth column shows the number of participants who did not paraphrase the sentence successfully, and the sixth column shows the percentage of participants who successfully paraphrased the sentence, which is calculated as the ratio of participants who paraphrased it successfully divided by

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

the total number of participants who did not repeat it verbatim. This percentage was treated as the sentence's difficulty index, and it was converted into a difficulty rank, which can be found in the column labeled "New diffic. rank," which does not take verbatim-repeated sentences into consideration. For purposes of comparison, the last two columns show the original difficulty rank and original item facility calculations for each sentence based on the original item scores in Pavlenko et al. (2019), which treat verbatim repetitions as instances of successful paraphrasing. Before turning to the primary results, we draw attention to the fact that the new and original difficulty ranks are nearly identical, suggesting that verbatim repetitions and successful paraphrases are generally indicative of the same ability.

Prediction 1a states that verbatim repetition will be more frequent among the more difficult sentences than among the easier ones. Our adopted threshold for confirming this prediction requires that the sentence with the most verbatim repetitions rank among the two most difficult sentences in the Pavlenko_WP dataset. The results provided in Table 2 do not confirm this prediction. The most difficult sentence (WP 4) was not repeated verbatim by any of the 100 participants, and only one participant (an L1 speaker) repeated the second most difficult sentence (WP 5). The only sentence that was repeated by more than two participants is WP 1, which is the second easiest sentence.

Corresponding results from the Akbary_OP dataset are shown in Table 3, which is set up in the same way as Table 2 except that Table 3 includes different stimulus sentences. As before, the question is whether verbatim repetition occurs more often with sentences that are more difficult to paraphrase successfully. Our adopted confirmation threshold for the Akbary_OP dataset requires that the sentence with the most verbatim repetitions rank among the 10 most difficult sentences. The results in Table 3 show that this is not the case. Even though every sentence was repeated verbatim by at least one of the 50 participants, repetition occurred less often with the more difficult sentences, and more often with the easier sentences, $r(28) = -.40$, $p = .029$, 95% CI[–.66, – .05]. The two sentences that were repeated by the largest number of participants were the third easiest (OP 2) and the second easiest (OP 4), respectively. Similar to what was mentioned in relation to Table 2, Table 3 shows that the new and original sentence difficulty ranks are very similar, $r(28) = .91$, $p < .001$, 95% CI [.82, .96].

To test Prediction 1b, we examined whether the participants with the lowest levels of sentence comprehension (i.e., the ones with the lowest rates of successful paraphrasing) were the most likely to resort to verbatim repetition. Our adopted confirmation criterion for this prediction requires that the participant with the most verbatim repetitions rank among the bottom third of participants in terms of the number of successful paraphrases that they produced. In the Pavlenko_WP dataset, we found that only four participants resorted to verbatim repetition in the Miranda paraphrasing task, three of whom were L1 participants. The one L2 participant who resorted to verbatim repetition repeated only one sentence (Sentence WP 1, the second easiest sentence). This L2 participant was ranked 55th out of the 100 participants (15th out of the 59 L2 participants) in terms of her comprehension index.

Corresponding results from our examination of the Akbary_OP dataset are shown in Table 4, where the 50 participants are listed in rank order (see the column labeled "New part. rank") from lowest to highest in terms of their comprehension index (see

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Table 2. Instances of repetition vs. paraphrasing, by sentence, in the Pavlenko_WP dataset

Sent. ID	Stimulus sentence	Raw verbat. repet.	Raw succ. paraph.	Raw unsuc. paraph.	% Succ. paraph.	New diffic. rank	Orig. diffic. rank	Orig. item facil.
WP 4	*If you cannot afford an attorney, one will be appointed to represent you before any questioning if you wish.*	0	7	93	.07	7	7	.32
WP 5	*You can decide at any time to exercise these rights and not answer any questions or make any statements.*	1	27	72	.27	6	6	.45
WP 3	*You have the right to talk to a lawyer and have him present with you while you are being questioned.*	2	31	67	.32	5	5	.53
WP 7	*Having these rights in mind do you wish to talk to us now?*	1	37	62	.37	4	4	.53
WP 2	*Anything you say can and will be used against you in a court of law.*	2	48	50	.49	3	2	.55
WP 1	*You have the right to remain silent.*	4	48	48	.50	2	3	.55
WP 6	*Do you understand each of these rights I have explained to you?*	2	58	40	.59	1	1	.68

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Table 3. Instances of repetition vs. paraphrasing, by sentence, in the Akbary_OP dataset

Sent. ID	Stimulus sentence	Raw verbat. repet.	Raw succ. paraph.	Raw unsuc. paraph.	% Succ. paraph.	New diffic. rank	Orig. diffic. rank	Orig. item facil.
OP 6	*I doubt that he knows how to drive that well.*	2	11	37	.23	30	29	0.48
OP 22	*The black cat that you fed yesterday was the one chased by the dog.*	3	15	32	.32	29	26	0.61
OP 16	*Cross the street at the light and then just continue straight ahead.*	1	16	33	.33	28	27	0.58
OP 29	*The exam wasn't nearly as difficult as you told me it would be.*	1	12	24	.33	27	28	0.57
OP 14	*You really enjoy listening to country music, don't you?*	4	16	30	.35	26	23	0.68
OP 25	*The terrible thief whom the police caught was very tall and thin.*	1	12	22	.35	25	30	0.45
OP 11	*The little boy whose kitten died yesterday is sad.*	8	9	15	.38	24	25	0.64
OP 20	*I hope it will get warmer sooner this year than it did last year.*	2	20	28	.42	23	21	0.69
OP 27	*The number of people who smoke cigars is increasing every year.*	7	17	23	.43	22	13	0.82
OP 26	*Would you be so kind as to hand me the book which is on the table?*	1	22	27	.45	21	19	0.75
OP 15	*She just finished painting the inside of her apartment.*	7	18	17	.51	20	20	0.74
OP 13	*I want a nice, big house in which my animals can live.*	3	17	16	.52	19	22	0.68
OP 21	*A good friend of mine always takes care of my neighbor's three children.*	2	23	20	.53	18	24	0.65
OP 12	*That restaurant is supposed to have very good food.*	9	22	19	.54	17	14	0.81

(*Continued*)

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Table 3. (*Continued.*)

Sent. ID	Stimulus sentence	Raw verbat. repet.	Raw succ. paraph.	Raw unsuc. paraph.	% Succ. paraph.	New diffic. rank	Orig. diffic. rank	Orig. item facil.
OP 24	*The most fun I've ever had was when we went to the opera.*	1	29	20	.59	16	16	0.79
OP 19	*I wish the price of town houses would become affordable.*	2	28	19	.60	15	17	0.78
OP 23	*Before he can go outside, he has to finish cleaning his room.*	4	33	21	.61	14	7	0.85
OP 28	*I don't know if the 11:30 train has left the station yet.*	2	30	18	.63	13	12	0.83
OP 8	*It is possible that it will rain tomorrow.*	9	26	15	.63	12	6	0.90
OP 17	*The person I'm dating has a wonderful sense of humor.*	2	34	18	.65	11	15	0.80
OP 18	*She only orders meat dishes and never eats vegetables.*	4	31	15	.67	10	10	0.85
OP 5	*What did you say you were doing today?*	4	27	12	.69	9	18	0.76
OP 7	*After dinner I had a long, peaceful nap.*	3	30	13	.70	8	11	0.83
OP 9	*I enjoy movies which have a happy ending.*	2	37	14	.73	7	5	0.90
OP 1	*I have to get a haircut.*	7	33	10	.77	6	7	0.85
OP 3	*The streets in this city are wide.*	1	37	11	.77	5	7	0.85
OP 30	*There are a lot of people who don't eat anything at all in the morning.*	4	36	10	.78	4	4	0.91
OP 2	*The red book is on the table.*	43	6	1	.86	3	1	0.99
OP 4	*He takes a shower every morning.*	17	29	4	.88	2	3	0.95
OP 10	*The houses are very nice but too expensive.*	8	39	3	.93	1	2	0.96

the column labeled "% Succ. paraph."). Our adopted confirmation criterion for this dataset requires that the participant with the most verbatim repetitions rank among the 17 lowest-performing participants in terms of successful paraphrasing. Contrary to Prediction 1b, Table 4 shows that participants with lower comprehension indices are less likely to repeat sentences verbatim, and participants with higher comprehension indices are more likely to repeat sentences, $r(48) = .50$, $p < .001$, 95% CI [–.68, –26]. The highest numbers of sentences repeated verbatim are seen among the L1 participants, with participants A046 and A047 repeating more than 20 of the 30 stimulus sentences.

The results compiled in Table 4 also allow us to address Prediction 1c, which states that the most difficult sentence that a given participant repeats will be more difficult than the most difficult sentence that the same person paraphrases successfully. The second to the last column in Table 4 shows the difficulty rank of the most difficult sentence that each participant successfully paraphrased, and the last column shows the difficulty rank of the most difficult sentence that the same participant repeated verbatim. Sentence ranks with lower numbers are easier, so Prediction 1c predicts that the number in the second to the last column will be smaller than the number in the last column of Table 4. This was found to be true for three of the 50 participants, two of whom are L1 participants (A046 and A047) who repeated more than 20 of the 30 sentences verbatim, including the most difficult sentence (rank 30). The one L2 participant who repeated a sentence that was more difficult than the most difficult one that she successfully paraphrased was Participant A026, who successfully paraphrased the third most difficult sentence (rank 28) and repeated the second most difficult sentence (rank 29). This participant was the second highest performing L2 participant and had a comprehension index higher than two of the L1 participants.

Our adopted confirmation criterion was that Prediction 1c is confirmed with any instance where the most difficult sentence that a participant has repeated verbatim exceeds the difficulty of the most difficult sentence that the same participant has paraphrased successfully, and where the difficulty ranks of the two sentences are not adjacent. This prediction is confirmed with the responses of the two L1 participants just mentioned (A046 and A047) but not by the response of the L2 participant (A026), whose most difficult repetition and paraphrase were adjacent in difficulty. In both cases that confirm the prediction, the repeated sentence was the most difficult sentence, Sentence OP 6, which consists of 10 syllables. Unlike the Akbary_OP dataset, the Pavlenko_WP dataset does not contain any instances where a participant repeated a sentence that was more difficult than one that they had paraphrased successfully.

To address Prediction 2, we compared the Akbary_OP and Akbary_EI datasets to determine whether participants were able to repeat sentences in the EI task that they could not successfully paraphrase in the paraphrasing task. This comparison involved 1,500 paired responses (50 participants × 30 sentences). In the EI data, 470 responses were found to be verbatim repetitions. The corresponding responses in the paraphrasing task broke down as follows: 150 verbatim repetitions, 286 successful paraphrases, and 34 unsuccessful paraphrases. Our adopted confirmation criterion for Prediction

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Table 4. Instances of repetition vs. paraphrasing, by participant, in the Akbary_OP dataset

Part. ID	L1 or L2	Raw verbat. repet.	Raw succ. paraph.	% Succ. paraph.	Part. total score	New part. rank	Orig. part. rank	Rank of most diff. sentence paraph. success.	Rank of most diff. sentence repeated verbatim
A017	L2	1	3	.10	21	50	50	20	3
A002	L2	0	4	.13	47	49	44	7	–
A021	L2	1	4	.14	35	48	49	19	3
A005	L2	1	5	.17	38	45	48	26	3
A010	L2	1	5	.17	40	45	46	12	3
A014	L2	1	5	.17	39	45	47	12	3
A009	L2	1	6	.21	46	44	45	22	3
A015	L2	1	7	.24	59	43	40	17	3
A020	L2	1	9	.31	57	42	42	14	3
A036	L2	0	10	.33	65	41	38	27	–
A035	L2	2	10	.36	64	40	39	29	3
A029	L2	1	11	.38	74	38	33	22	3
A037	L2	1	11	.38	72	38	35	21	3
A003	L2	1	12	.41	55	35	43	20	6
A004	L2	1	12	.41	58	35	41	15	3
A018	L2	1	12	.41	75	35	32	28	3
A007	L2	2	12	.43	69	31	36	28	3

(*Continued*)

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Table 4. (*Continued.*)

Part. ID	L1 or L2	Raw verbat. repet.	Raw succ. paraph.	% Succ. paraph.	Part. total score	New part. rank	Orig. part. rank	Rank of most diff. sentence paraph. success.	Rank of most diff. sentence repeated verbatim
A027	L2	2	12	.43	80	31	28	23	3
A032	L2	2	12	.43	68	31	37	28	5
A039	L2	2	12	.43	77	31	30	21	9
A044	L2	1	13	.45	81	30	26	23	3
A013	L2	2	14	.50	73	29	34	22	3
A011	L2	1	15	.52	77	27	30	26	3
A034	L2	1	15	.52	81	27	26	23	3
A045	L2	5	13	.52	84	26	23	29	19
A025	L2	2	15	.54	82	25	25	21	3
A006	L2	1	16	.55	78	23	29	27	3
A028	L2	1	16	.55	84	23	23	29	3
A033	L2	0	18	.60	88	22	21	29	–
A043	L2	2	17	.61	88	21	21	30	3
A030	L2	2	18	.64	89	20	20	30	22
A023	L2	2	19	.68	95	19	14	29	20
A042	L2	8	15	.68	93	18	18	30	26

(*Continued*)

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Table 4. (*Continued.*)

Part. ID	L1 or L2	Raw verbat. repet.	Raw succ. paraph.	% Succ. paraph.	Part. total score	New part. rank	Orig. part. rank	Rank of most diff. sentence paraph. success.	Rank of most diff. sentence repeated verbatim
A038	L2	6	17	.71	94	17	15	28	24
A012	L2	1	21	.72	93	16	18	26	3
A024	L2	4	19	.73	94	15	15	25	12
A040	L2	1	22	.76	96	14	13	30	3
A041	L2	3	21	.78	98	13	12	29	20
A022	L2	0	24	.80	94	12	15	28	–
A048	L1	8	18	.82	101	11	10	30	24
A046	L1	24	5	.83	105	10	5	28	30
A016	L2	4	22	.85	100	9	11	30	22
A026	L2	8	19	.86	102	8	8	28	29
A031	L2	4	23	.88	102	7	8	30	20
A051	L1	9	19	.90	104	6	6	29	24
A049	L1	0	28	.93	103	5	7	30	–
A047	L1	26	4	1.00	106	1	1	22	30
A050	L1	10	20	1.00	106	1	1	30	24
A052	L1	5	25	1.00	106	1	1	30	24
A053	L1	0	30	1.00	106	1	1	30	–

2 is the occurrence of any instance where a participant has produced a verbatim repetition of the stimulus sentence in the EI task but has failed to produce a successful paraphrase of the same stimulus sentence in the paraphrasing task. Our analysis thus shows that Prediction 2 is confirmed by the 34 cases where a participant repeated a sentence in the EI task but failed to paraphrase it successfully in the paraphrasing task.

To give consideration to the totality of the circumstances, we examined the 34 cases of unsuccessful paraphrasing in the Akbary_OP dataset that correspond with verbatim repetitions in the Akbary_EI dataset. The results of this additional analysis are summarized in Table 5, where stimulus sentences are ordered from most difficult to least difficult in accordance with the rankings provided earlier in Table 3. Unlike Table 3, which shows all 30 stimulus sentences, Table 5 lists only the 20 stimulus sentences that were repeated verbatim in the EI task and paraphrased unsuccessfully in the oral paraphrasing task. The first two columns of Table 5 list the relevant stimulus sentences and the third column shows the difficulty ranks for these sentences (corresponding with the ranks shown in the "New diffic. rank" column in Table 3). The fourth column in Table 5 shows the length of the stimulus sentence in syllables, and the fifth column shows the number of participants who repeated this sentence verbatim in the EI task but failed to paraphrase it successfully in the paraphrasing task. The final two columns list the participants who failed to paraphrase these sentences successfully. As described earlier in our *Analysis* section, cases of failed paraphrasing can provide relatively weaker or relatively stronger evidence of noncomprehension. Weaker, more ambiguous cases include unsuccessful paraphrases that are incomplete but semantically aligned with the stimulus sentence. Stronger and more direct evidence of noncomprehension includes responses that contradict the semantic content of the stimulus sentence. The participants whose responses constitute weaker evidence of noncomprehension are listed in Table 5 in the column labeled "Weak evidence (participant ID)," whereas the participants whose responses represent stronger evidence of noncomprehension are listed in the column labeled "Strong evidence (participant ID)." Examples of each type are illustrated in Examples 1 and 2 below:

Example 1 (weak evidence): Participant A006's responses to Sentence EI 6/OP 6
EI response: *I doubt that he knows how to drive that well.*
Oral paraphrase response: *He say that I don't know how to drive.*

Example 2 (strong evidence): Participant A026's responses to Sentence EI 6/OP 6
EI response: *I doubt that he knows how to drive that well.*
Oral paraphrase response: *He said that he doubts that he knows how to draw that well.*

In Example 1, Participant A006's paraphrase of the stimulus sentence appears to be an attempt at reported speech, which was the typical pattern in the Akbary_OP dataset. In this example, the participant might have understood the speaker in the audio recording to be referring to the participant's lack of driving ability, which would account for why "he" in the stimulus sentence became "I" in the oral paraphrase response. If this is true, then Participant A006's paraphrase in Example 1 could represent an instance of successful comprehension even though it was rated as an unsuccessful

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

paraphrase. Example 2, on the other hand, unambiguously shows that the participant misunderstood the sentence by misinterpreting the word "drive" as "draw."

Table 5 shows that the 34 instances of unsuccessful paraphrases that confirm Prediction 2 involve 20 stimulus sentences ranking in difficulty from the most difficult (i.e., rank 30 in the column labeled "Diff. rank") to the second easiest (rank 2). Sentence lengths vary from 7 to 18 syllables. The 34 cases of unsuccessful paraphrasing were produced by 22 participants. Two participants (A016 and A024) produced four unsuccessful paraphrases apiece, one additional participant (A024) produced three instances, and another five participants produced two apiece. The 12 cases representing strong, unambiguous cases of miscomprehension were produced by eight participants. One of these participants (A016) produced four unambiguous instances of miscomprehension, and one additional participant (A026) produced two.

Although the 34 instances of failed paraphrase confirm Prediction 2, and at least 12 of these instances (i.e., those listed in the final column of Table 5) unambiguously demonstrate the participant's failure to comprehend the stimulus sentence, we recognize that it is possible for participants to understand a sentence during one task and fail to comprehend the same sentence during a different task administered on a different day due to fluctuating task demands, context, and environmental conditions that interact with cognitive resources (Vaughan & Birney, 2023). Since the core question in this analysis is whether the verbatim repetitions in the EI task were rote repetitions produced without comprehension, and because we do not have direct evidence of whether the participants understood these sentences at the time of the EI administration, it is useful to consider whether the participants' patterns of successful paraphrasing suggest that the sentences they repeated verbatim were likely to be beyond their level of comprehension. Of all 22 participants listed in the final two columns of Table 5, only five repeated a sentence on the EI task that was more difficult than any of the sentences that they successfully paraphrased on the oral paraphrasing task. Additionally, two of these five participants produced a verbatim repetition on the EI task that differed by only one rank from the most difficult sentence they successfully paraphrased on the paraphrasing task (Participant A024 successfully paraphrased the sentence with difficulty rank 25 and repeated the sentence with difficulty rank 26; Participant A051 successfully paraphrased difficulty rank 29 and repeated difficulty rank 30). This leaves just three instances where a participant repeated a sentence on the EI task that had a difficulty rank more than one rank higher than the most difficult sentence that they paraphrased successfully: Participant A026 repeated the most difficult sentence (OP 6, rank 30) but did not demonstrate an ability to successfully paraphrase a sentence above rank 28; Participant A006 repeated the most difficult sentence (OP 6, rank 30) but was not able to successfully paraphrase beyond difficulty rank 27; and Participant A025 repeated the fifth most difficult sentence (OP 14, rank 26) but was not able to paraphrase successfully beyond rank 21 on the paraphrasing task.

However, the last participant (A025) did successfully paraphrase the third most difficult sentence (OP 16) on the EI task, suggesting that the fifth most difficult sentence was not beyond this person's comprehension level. Additionally, Participant A006's paraphrase of Sentence OP 6 – while considered unsuccessful by the raters in Akbary et al. (2023) – is the same sentence illustrated earlier in Example 1, which may reflect

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Table 5. Instances of unsuccessful paraphrasing in the Akbary_OP dataset that correspond with verbatim repetitions in the Akbary_EI dataset

Sent. ID	Stimulus sentence	Diffic. rank	Sent. length	Unsuc. paraph.	Weak evidence (participant ID)	Strong evidence (participant ID)
OP 6	*I doubt that he knows how to drive that well.*	30	10	3	A006, A051	A026
OP 22	*The black cat that you fed yesterday was the one chased by the dog.*	29	16	1		A016
OP 16	*Cross the street at the light and then just continue straight ahead.*	28	15	1		A048
OP 29	*The exam wasn't nearly as difficult as you told me it would be.*	27	18	1		A016
OP 14	*You really enjoy listening to country music, don't you?*	26	14	4	A024, A025, A031, A033	
OP 25	*The terrible thief whom the police caught was very tall and thin.*	25	17	2	A049	A051
OP 11	*The little boy whose kitten died yesterday is sad.*	24	13	2	A027	A012
OP 27	*The number of people who smoke cigars is increasing every year.*	22	17	1		A027
OP 26	*Would you be so kind as to hand me the book which is on the table?*	21	17	1	A048	
OP 12	*That restaurant is supposed to have very good food.*	17	13	1	A011	

(*Continued*)

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Table 5. (*Continued.*)

Sent. ID	Stimulus sentence	Diffic. rank	Sent. length	Unsuc. paraph.	Weak evidence (participant ID)	Strong evidence (participant ID)
OP 24	*The most fun I've ever had was when we went to the opera.*	16	16	1		A026
OP 8	*It is possible that it will rain tomorrow.*	12	12	4	A027, A035, A043, A044	
OP 17	*The person I'm dating has a wonderful sense of humor.*	11	15	1	A027	
OP 18	*She only orders meat dishes and never eats vegetables.*	10	15	1		A016
OP 5	*What did you say you were doing today?*	9	10	3	A024, A034, A042	
OP 7	*After dinner I had a long, peaceful nap.*	8	11	1	A039	
OP 9	*I enjoy movies which have a happy ending.*	7	12	1	A023	
OP 1	*I have to get a haircut.*	6	7	1	A039	
OP 2	*The red book is on the table.*	3	8	1	A002	
OP 4	*He takes a shower every morning.*	2	9	3		A016, A021, A033

successful comprehension. An analysis considering a totality of the circumstances therefore points to Participant A026's performance on Sentence OP 6 as the strongest evidence for repetition without comprehension. Participant A026 is the same participant mentioned earlier in relation to Prediction 1c (Table 4), who was the only L2 participant to repeat a sentence on the oral paraphrasing task (Sentence OP 22, rank = 29) that was more difficult than a sentence that she successfully paraphrased on the same task (Sentence OP 16, rank = 28). Her performance on the EI task, which she took the day before the oral paraphrasing task, shows that she repeated 16 of the 30 stimulus sentences verbatim, including all three of the most difficult sentences.

Discussion

The results of the present study did not confirm Predictions 1a or 1b. Our results showed that the sentences that were more difficult to comprehend (i.e., to paraphrase successfully) were not repeated verbatim more than sentences that were easier to comprehend (Prediction 1a), and that participants with greater comprehension challenges did not repeat sentences verbatim more than higher-performing participants (Prediction 1b). These predictions deal with generalizable trends rather than specific instances, and our results are perhaps not surprising given that they mirror those of past quantitative research that has similarly failed to find patterns of rote repetition occurring with sentences longer than 10 syllables (e.g., Deygers, 2020; Vinther, 2002).

Our analyses related to Predictions 1c and 2 more directly dealt with the question of whether even a single instance of rote repetition without comprehension can be found. Both predictions were confirmed. Regarding Prediction 1c, we found two instances where the most difficult sentence that a participant repeated verbatim on the oral paraphrasing task (Akbary_OP dataset) was more difficult than the most difficult sentence that the same person paraphrased successfully in the same task. However, the two participants who produced these instances were the two L1 participants who resorted to verbatim repetition as their default strategy when performing the oral paraphrasing task. One of these two participants repeated verbatim 24 of the 30 stimulus sentences, and the other participant repeated 26 of these sentences (see Table 4). It appears that the reason that these two participants did not successfully paraphrase a sentence that was more difficult than the ones they repeated verbatim is because they repeated most sentences verbatim – including the most difficult ones. Even though Prediction 1c was confirmed, the details of the data cast doubt on whether they truly reflect rote repetition without comprehension.

Prediction 2 was likewise confirmed but with some important caveats. Our analysis for Prediction 2 involved a comparison of pairs of responses in the Akbary_EI and Akbary_OP datasets. We found that 34 (7.23%) of the 470 verbatim repetitions in the EI task corresponded to unsuccessful paraphrases in the paraphrasing task. If the 22 participants who produced these unsuccessful paraphrases truly did understand these sentences while performing the EI task, then why were they not able to successfully paraphrase the same sentences on the paraphrasing task? There could be several reasons, including the greater cognitive load imposed by the paraphrasing task (Khrismawan & Widiati, 2013), the fact that paraphrasing tasks often result in incomplete restatements of the stimulus (Wolfe & Goldman, 2005), and the possibility that

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

the external or participant-internal test-taking conditions were less optimal during the paraphrasing task than during the EI task for these 22 participants (Shang et al., 2024). In any case, we found that only three of the 22 participants repeated a sentence on the EI task that was more difficult than any of the sentences that they successfully paraphrased on the oral paraphrasing task. One of these three participants did successfully paraphrase a sentence on the EI task that was more difficult than the most difficult one that they repeated verbatim on the same task. Another of the three participants produced a paraphrase that – although it was rated as unsuccessful by the raters in the original Akbary et al. (2023) study – might reflect successful comprehension (see Example 1). The most compelling instance of rote repetition without comprehension is Participant A026's exact repetition of the most difficult sentence (OP 6) on the EI task and her clear miscomprehension of this same sentence on the paraphrasing task (see Example 2). There is little question that this participant misunderstood the sentence while performing the paraphrasing task, but the idea that this sentence was too difficult for her to comprehend is called into question by the fact that she was one of the three highest-performing L2 participants on both the EI and paraphrasing task regardless of how performance is measured (see Table 4). Her demonstrated high level of performance combined with the fact that she successfully paraphrased the third most difficult sentence (Sentence OP 16) on the paraphrasing task suggests that her verbatim repetitions might not reflect a lack of sentence comprehension.

The literature we reviewed on nonword repetition and related phenomena suggests that verbatim repetition without comprehension is typical for strings of up to four syllables, but becomes increasingly improbable from five syllables onward, and becomes essentially impossible beyond nine or ten syllables (Deygers, 2020; Gathercole et al., 1991, 1994; Graf-Estes et al., 2007; Gupta & MacWhinney, 1997; Vinther, 2002). The sentence lengths for the Miranda sentences in the Pavlenko_WP dataset range from 9 to 31 syllables, and the sentence lengths for the 30 sentences in Akbary_OP range from 7 to 19. Sentences in both paraphrasing datasets consisting of fewer than 10 syllables were indeed repeated the most (Sentence WP 1 in Table 3 consists of 9 syllables; Sentences OP 2 and OP 4 in Table 4 consist of 8 and 9 syllables, respectively). We acknowledge that some of the participants who repeated these sentences verbatim might have relied on phonological short-term memory rather than on semantic memory (cf. Horne et al., 2022). However, this interpretation is somewhat dubious given that the other 2 sentences consisting of fewer than 10 syllables (OP 1 with 7 syllables, OP 3 with 8 syllables) do not show the same high levels of verbatim repetition. In any event, we note that if Participant A026's performance on the most difficult sentence (OP 6, rank 30) is indeed an instance of rote repetition without comprehension, this sentence – which consists of 10 syllables – is at the upper limit of what the relevant literature suggests is possible for repetition without comprehension (e.g., Deygers, 2020).

Other reasons might offer better explanations for why participants sometimes repeat sentences verbatim even when they are asked to restate them in their own words. The authors of the present study recently tested this informally with MA students at an American university. The stimulus sentences from Akbary_OP were played to the students, who were asked to restate each sentence in their own words. The students' responses were similar to those of the Akbary_OP dataset. When asked why they repeated some of the sentences even though they were asked to paraphrase them,

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

several of the students indicated that they did so because there was no other immediately obvious way to express the same meaning with different words. We suspect that this explanation might account for why certain sentences such as OP 2 and OP 4 were repeated so often in the oral paraphrasing task (see Table 3).

Although Prediction 1c was confirmed with two instances of potential rote repetition without comprehension, and Prediction 2 was confirmed with 34, we do not believe that a judge, jury, or panel of linguists would find that any of these 36 instances, when interpreted in light of the totality of the available evidence, renders it more likely than not that repetition without comprehension has occurred. At the same time, we believe that our results are sufficiently indicative of the possibility of sentence repetition without comprehension to warrant further research, as we elaborate in the following section.

Regarding the purpose and context of the present study, and to the extent that the data examined are relevant to the assessment of Miranda comprehension by forensic practitioners, we find that verbatim repetitions – whether produced during an EI or paraphrasing task – are overwhelmingly in alignment with patterns of successful paraphrasing and thus also patterns of sentence comprehension. However, as we discussed in the section titled *Sentence paraphrasing as a measure of listening comprehension*, a successful paraphrase provides more compelling evidence of comprehension than does successful repetition given that producing a perfect paraphrase requires reformulation without loss of meaning, which is impossible without comprehension.

For this reason, we believe that Recommendation 6 of the CoRG guidelines is well-formulated for its intended audience – i.e., police officers who deliver the Miranda warning to arrestees. Because they must ensure that the arrestee comprehends each statement, it makes sense for law enforcement to try to elicit a perfect paraphrase of each. Forensic practitioners, conversely, seek to identify instances of noncomprehension. These are often less ambiguous in EI tasks than in paraphrasing tasks (see Akbary et al., 2023) given that, in EI tasks but not paraphrasing tasks, the participant can reasonably be assumed to have been trying to reproduce every element of meaning from the stimulus sentence. While underscoring the importance of failed repetitions as indicators of noncomprehension, the present study explored whether even successful repetitions might occasionally occur in the absence of comprehension. We found that verbatim repetitions without comprehension occur somewhere between 0% and 7.23% of the time, with the totality of the evidence favoring the lower end of this range. We discuss implications for future research in the following section.

Limitations and future directions

The present study has important limitations related primarily to the fact that we relied on data from previously published studies that were not designed to test whether sentence repetition is possible without comprehension. The datasets from Akbary et al. (2023) also lacked direct relevance to Recommendation 6 from the CoRG guidelines in that they did not include any Miranda sentences. The dataset from Pavlenko et al. (2019) did consist of Miranda sentences, but its use of written instead of oral responses differs from the CoRG guidelines and from the manner in which Miranda comprehension is typically assessed. Perhaps the most serious limitation of the datasets from both

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

studies is that they lacked a second source of evidence for listening comprehension when the participant repeated a sentence instead of paraphrasing it. The two datasets from Akbary et al. did address this shortcoming to a certain degree by including both an EI task and a paraphrasing task, where participants' responses on the paraphrasing task often demonstrated whether they comprehended the sentences they repeated verbatim on the EI task. However, the fact that the two tasks were administered at different times leaves open the possibility that participants could have comprehended a sentence during one of the tasks and failed to comprehend the same sentence during the other task (e.g., Shang et al., 2024; Vaughan & Birney, 2023).

A future study could overcome these shortcomings and determine with greater levels of confidence whether the types of verbatim repetition that confirmed – with caveats – two of our predictions are indeed instances of failed comprehension. Most importantly, such a study would include an additional measure of comprehension immediately after each verbatim repetition to determine whether the participant did or did not understand that specific sentence on that specific occasion. To avoid the limitations of multiple-choice items and paraphrasing tasks, we believe that it would be valuable to have participants either orally translate the sentence into their L1 or engage in a verbal protocol in which they verbalize their thought processes, understanding of individual words and phrases, and the strategies they choose while recalling the sentence a second time (perhaps without hearing it a second time). The use of a verbal protocol might also elucidate why participants sometimes resort to verbatim repetitions when they have been asked to paraphrase. The use of either verbal protocols or L1 translation might also shed additional light on the effects of productive skills on repetition failures and unsuccessful paraphrases.

Conclusion

The present study analyzed sentence paraphrasing and sentence repetition data from two previous studies to test four predictions related to the question of whether an individual can repeat a sentence verbatim without understanding it. Two predictions addressing this question were confirmed. However, a broader analysis of the instances of verbatim repetition that confirmed these predictions suggested that they might not have reflected failed comprehension. Further research will be necessary to resolve the ambiguity of our findings, and we have offered some suggestions for how to do this.

Regardless of whether individuals can repeat Miranda warning sentences without comprehending them, we recognize the value of Recommendation 6 of the CoRG guidelines for the purposes it was created. It is a recommendation for law enforcement officials responsible for delivering the Miranda warning to detainees and ensuring that they have understood it fully. Given that paraphrasing provides stronger evidence of full comprehension than verbatim repetition does, Recommendation 6 errs on the side of caution in an attempt to ensure that detainees truly do understand their rights and have the opportunity to exercise them before the police interrogation begins. However, it is important to recognize that Recommendation 6 is not a field standard for forensic linguists and should not be cited as such in one's expert testimony. Because the assessment of Miranda comprehension by forensic experts focuses on evidence of the lack of comprehension, linguists who engage in this type of work need to recognize the

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

value of EI as a method for eliciting relatively unambiguous evidence of failed comprehension: Would a person who understood this sentence have produced such a flawed repetition of it?

References

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, *135*(1), 50–68. https://doi.org/10.1037/a0014411

Akbary, M., Benzaia, L. A., Jarvis, S., & Park, H. I. (2023). Evaluating the utility of elicited imitation as a measure of listening comprehension and Miranda competency. *Research Methods in Applied Linguistics*, *2*(3), 100067. https://doi.org/10.1016/j.rmal.2023.100067

Aryadoust, V., & Luo, L. (2023). The typology of second language listening constructs: A systematic review. *Language Testing*, *40*(2), 375–409. https://doi.org/10.1177/02655322221126604

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. https://doi.org/10.1016/S0022-5371(75)80045-4

Benzaia, L. A., Jarvis, S., Akbary, M., & Park, H. I. (2024). Investigating the relationship between L2 proficiency and Miranda rights comprehension: A partial replication of Pavlenko et al. (2019). *International Journal of Speech, Language and the Law*, *31*(2), 291–327. https://doi.org/10.3138/ijsll-2024-0018

Bley-Vroman, R., & Chaudron, C. (1994). Elicited imitation as a measure of second-language competence. In E. E. Tarone, S. M. Gass & A. D. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 245–261). Erlbaum.

Bonk, W. J. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, *14*(1), 14–31. https://doi.org/10.1080/10904018.2000.10499033

Bowden, H. W. (2016). Assessing second-language oral proficiency for research: The Spanish elicited imitation task. *Studies in Second Language Acquisition*, *38*(4), 647–675. https://doi.org/10.1017/S0272263115000443

Buck, G. (2001). *Assessing listening*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732959

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. https://doi.org/10.1037/h0046016

Communication of Rights Group (CoRG) (2015). *Guidelines for communicating rights to nonnative speakers of English in Australia, England and Wales, and the USA*. Retrieved from November 21, 2025, https://www.aaal.org/guidelines-for-communication-rights.

Cox, T. L., & Davies, R. S. (2012). Using automatic speech recognition technology with elicited oral response testing. *CALICO Journal*, *29*(4), 601–618. https://doi.org/10.11139/cj.29.4.601-618

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

Davis v. United States, 512 U.S. 452. (1994).

Deygers, B. (2020). Elicited imitation: A test for all learners? Examining the EI performance of learners with diverging educational backgrounds. *Studies in Second Language Acquisition*, *42*(5), 933–957. https://doi.org/10.1017/S027226312000008X

Ebert, K. D. (2014). The role of auditory nonverbal working memory in sentence repetition for bilingual children with primary language impairment. *International Journal of Language and Communication Disorders*, *49*(5), 631–636. https://doi.org/10.1111/1460-6984.12090

Eggington, W. G., & Cox, T. (2013). Using elicited oral response testing to determine the need for an interpreter. *Harvard Latino Law Review*, *16*, 127–146. https://www.academia.edu/6084999/USING_ELICITED_ORAL_RESPONSE_TESTING_TO_DETERMINE_THE_NEED_FOR_AN_INTERPRETER

Fare v. Michael C., 442 U.S. 707 (1979).

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Frizelle, P., O'Neill, C., & Bishop, D. V. M. (2017). Assessing understanding of relative clauses: A comparison of multiple-choice comprehension versus sentence repetition. *Journal of Child Language*, *44*(6), 1435–1457. https://doi.org/10.1017/S0305000916000635

Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics*, *12*, 349–367. https://doi.org/10.1017/S0142716400009267

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*(2), 103–127. https://doi.org/10.1080/09658219408258940

Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, *28*(1), 55–75. https://doi.org/10.1016/S0346-251X(99)00060-3

Goh, C. C. M., & Vandergrift, L. (2021). *Teaching and learning second language listening: Metacognition in action* (2nd ed.). Routledge. https://doi.org/10.4324/9780429287749

Graf-Estes, K., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *50*(1), 177–195. https://doi.org/10.1044/1092-4388(2007/015)

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language*, *59*(2), 267–333. https://doi.org/10.1006/brln.1997.1819

Hogan, T. P., Adlof, S. M., & Alonzo, C. (2014). On the importance of listening comprehension. *International Journal of Speech-Language Pathology*, *16*(3), 199–207. https://doi.org/10.3109/17549507.2014.904441

Horne, A., Zahn, R., Najera, O. I., & Martin, R. C. (2022). Semantic working memory predicts sentence comprehension performance: A case series approach. *Frontiers in Psychology*, *13*, 887586. https://doi.org/10.3389/fpsyg.2022.887586

Khrismawan, B., & Widiati, U. (2013). Students' perceptions about paraphrasing and their cognitive processes in paraphrasing. *TEFLIN Journal*, *24*(2), 136–157. https://journal.teflin.org/index.php/journal/article/view/162

Kim, Y., Tracy-Ventura, N., & Jung, Y. (2016). A measure of proficiency or short-term memory? Validation of an elicited imitation test for SLA research. *Modern Language Journal*, *100*(3), 655–673. https://doi.org/10.1111/modl.12346

Kostromitina, M., & Plonsky, L. (2022). Elicited imitation tasks as a measure of L2 proficiency: A meta-analysis. *Studies in Second Language Acquisition*, *44*(4), 886–911. https://doi.org/10.1017/S0272263121000395

McMenamin, G. H. (2022). Mourning the slow death of Miranda: California v. Ceja. In I. Picornell, R. Perkins & M. Coulthard (Eds.), *Methodologies and challenges in forensic linguistic casework* (pp. 95–113). Wiley. https://doi.org/10.1002/9781394266661.ch7

Meng, Y., Wang, Y., & Zhao, N. (2023). Cognitive diagnostic assessment of EFL learners' listening barriers through incorrect responses. *Frontiers in Psychology*, *14*, 1126106. https://doi.org/10.3389/fpsyg.2023.1126106

Miranda v. Arizona, 384 U.S. 436 (1966).

Mirzaei, M. S., & Meshgi, K. (2020). Sentence complexity as an indicator of L2 learner's listening difficulty. In K.-M. Frederiksen, S. Larsen, L. Bradley & S. Thouësny (Eds.), *CALL for widening participation: Short papers from EUROCALL 2020* (pp. 227–232). Research–publishing.net. https://doi.org/10.14705/rpnet.2020.48.1193

Naiman, N. (1974). The use of elicited imitation in second language acquisition research. *Working Papers on Bilingualism*, *2*, 1–37. https://www.academia.edu/687563/Elicited_imitation_in_second_language_acquisition_research.

Okura, E., & Lonsdale, D. (2012). Working memory's meager involvement in sentence repetition tests. In N. Miyake, D. Peebles & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 2132–2137). Cognitive Science Society.

Ortega, L., Iwashita, N., Norris, J. M., & Rabie, S. (2002). *An investigation of elicited imitation tasks in crosslinguistic SLA research*. Paper presented at the Second Language Research Forum.

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press

Park, H. I., Solon, M., Henderson, C., & Dehghanchaleshtori, M. (2020). The roles of working memory and oral language abilities in elicited imitation performance. *The Modern Language Journal*, *104*(1), 133–151. https://doi.org/10.1111/modl.12618

Pavlenko, A. (2008). 'I'm very not about the law part': Nonnative speakers of English and the Miranda warnings. *TESOL Quarterly*, *42*(1), 1–30. https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1545-7249.2008.tb00205.x

Pavlenko, A., Hepford, E., & Jarvis, S. (2019). An illusion of understanding: How native and non-native speakers of English understand (and misunderstand) their Miranda rights. *International Journal of Speech Language and the Law*, *26*(2), 181–207. https://doi.org/10.1558/ijsll.39163

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language*, *29*(6), 633–654. https://doi.org/10.1016/0749-596X(90)90042-X

Rogers, R., & Drogin, E. Y. (2019). *Conducting Miranda evaluations: Applications of psychological expertise and science within the forensic context*. Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-13511-9

Rujas, I., Mariscal, S., Murillo, E., & Lázaro, M. (2021). Sentence repetition tasks to detect and prevent language difficulties: A scoping review. *Children*, *8*(7), 578. https://doi.org/10.3390/children8070578

Shang, Y., Aryadoust, V., & Hou, Z. (2024). A meta-analysis of the reliability of second language listening tests (1991-2022). *Brain Sciences*, *14*(8), 746. https://doi.org/10.3390/brainsci14080746

Vaughan, A. C., & Birney, D. P. (2023). Within-individual variation in cognitive performance is not noise: Why and how cognitive assessments should examine within-person performance. *Journal of Intelligence*, *11*(6), 110. https://doi.org/10.3390/jintelligence11060110

Vinther, T. (2002). Elicited imitation: A brief overview. *International Journal of Applied Linguistics*, *12*(1), 54–73. https://doi.org/10.1111/1473-4192.00024

West, D. (2012). Elicited imitation as a measure of morphemic accuracy: Evidence from L2 Spanish. *Language and Cognition*, *4*(3), 203–222. https://doi.org/10.1515/langcog-2012-0011

Wolfe, M. B., & Goldman, S. R. (2005). Relations between adolescents' text processing and reasoning. *Cognition and Instruction*, *23*(4), 467–502. https://doi.org/10.1207/s1532690xci2304_2

Wu, S.-L., Tio, Y. P., & Ortega, L. (2022). Elicited imitation as a measure of L2 proficiency: New insights from a comparison of two L2 English parallel forms. *Studies in Second Language Acquisition*, *44*(2), 271–300. https://doi.org/10.1017/S0272263121000103

Cite this article: Jarvis, S., & Akbary, M. (2026). Sentence repetition, paraphrasing, and comprehension: Can you repeat a sentence without understanding it? *Annual Review of Applied Linguistics*, *46*, 144–178. https://doi.org/10.1017/S0267190526100269

https://doi.org/10.1017/S0267190526100269 Published online by Cambridge University Press