Hesson A, Pichler H. Interpreting "I don’t know" use by persons living with dementia in Mini-Mental State Examinations. *Patient Education and Counseling* 2016, 99(9), 1534-1541.

DOI link to article:

http://dx.doi.org/10.1016/j.pec.2016.07.001

Date deposited:

05/09/2016

Embargo release date:

02 July 2017

Interpreting "I don't know" Use by Persons Living with Dementia in Mini-Mental State Examinations

Ashley M. Hesson, **M.D., Ph.D. (corresponding author)**

Department of Obstetrics and Gynecology

University of Michigan

1500 E. Medical Center Drive

L3211 Women's, SPC 5276

Ann Arbor, MI 48109-5276

ahesson@med.umich.edu

+1 (716) 348-9693

Heike Pichler, Ph.D.

School of English Literature, Language & Linguistics

Percy Building

Newcastle University

Newcastle upon Tyne, NE1 7RU, UK

heike.pichler@newcastle.ac.uk

+44 (778) 663-2748

Conflicts of Interest: The authors have no conflicts of interest to declare.

Abstract word count: 197

Main text word count: 4541

Number of tables: 5

Number of figures: 0

Interpreting "I don't know" Use by Persons Living with Dementia in Mini-Mental State Examinations

ABSTRACT

Objective: We investigate dementia patients' use of I DON'T KNOW (IDK) in Mini-Mental State Exams (MMSEs) using objective linguistic indicators to differentiate IDK signalling lack of knowledge (LOK) from IDK used to hedge responses, affect exam progression etc. We hypothesize that increased proportional use of LOK-IDK correlates with worsening dementia severity.

Methods: 189 IDK tokens were extracted from 72 MMSE interactions and coded for linguistic/social characteristics. A data-driven, discourse position/relation-based functional taxonomy for IDK in MMSE was developed and the resulting functional distribution was subjected to multiple logistic regression.

Results: Use of LOK-IDK (vs. non-LOK-IDK) is significantly correlated (p=0.01) with clinicians' subjective ratings of patients' dementia as 'severe' vs. 'mild'/'moderate', indicating that objective sociolinguistic criteria approximate physician judgments. 92% of 'severe' patients' IDKs signalled LOK, compared to only 68% of 'mild' patients', suggesting that uncritical interpretation of IDK as signalling LOK would result in 8-32% of IDK responses being mis-scored.

Conclusion: LOK and non-LOK uses distinguished on the basis of reliable, objective usage patterns are differentially distributed among dementia severity groups.

Practice implications: LOK-IDK serves as a supplemental indicator of dementia severity. Correct interpretation may improve diagnostic accuracy and allow clinicians to respond supportively during cognitive assessment.

Keywords: sociolinguistics; discourse-pragmatic; severity; Alzheimer's; cognitive assessment

1. Introduction

Clinicians report uncertainty in differentiating pathological, progressive **cognitive impairment (i.e. dementia)** from cognitive decline associated with normal aging, functional memory disorders, and depression [1,2,3]. This lack of confidence may stem from providers' discomfort with explicit cognitive assessment, a potentially embarrassing [2] or outright confrontational [4] process that is face-threatening to patients [5,6] and often exposes stigmatized cognitive deficits [7]. However, few dementia-specific, evidence-based communication curricula exist to assist providers in the negotiation of communicative barriers to productive cognitive assessment [8,9].

Recent work suggests that subtle differences in the way patients respond to conversational prompts may provide reliable cues to the presence and severity of **cognitive impairment**. Specifically, impaired patients' atypical use of "I don't know"—a multifunctional phrase that can be used to disavow knowledge, signal speaker stance and affect discourse organization [10]— may hint at underlying cognitive dysfunction. Mikesell's [11] conversation analysis of a frontotemporal dementia patient's routine interactions in non-clinical environments contrasts this individual's appropriate and problematic uses of "I don't know" (henceforth IDK) as a response to WH-questions **(i.e., interrogatives constructed with "who", "why", "how", etc.)**. The problematic uses occur in response to questions that have clearly accessible answers, either based on the subject's current activity (e.g., "What are you reading?"), his close relations (e.g., "What does your daughter do?"), or his prior demonstrations of knowledge. They are inconsistent with the conversational task at-hand and frustrate the natural progression of discourse, similar to resistive IDK uses in child mental health consultations [12,13]. In the context of cognitive impairment, however, Mikesell claims that her subject is unlikely to be intentionally resisting a line of questioning. Based on the frequency of his inappropriate IDK uses and his interlocutors' treatment of these uses (usually involving elaborate repair sequences), she asserts that their occurrence should be taken as evidence of the subject's impairment.

Elsey and colleagues [14] offer additional support for IDK's utility in assessing **cognitive impairment** through their study of its production in memory clinic interactions in the UK. Focusing on the open-ended assessment portion of the visits, they find that IDK production frequency is significantly higher in those with **cognitive impairment** as compared to those with deficits attributable to functional memory disorders. However, Elsey and co-workers did not undertake a functional analysis of IDK. They treated all IDKs as interchangeable, even though linguists have repeatedly shown IDK to have two macro-roles: 1) as a claim to a cognitive state (= lack of knowledge), and 2) as a non-cognitive interactional device (= resistance strategy, epistemic marker, turn-exchange signal etc.) [15,12,13,16,17]. Furthermore, by excluding the formal assessment segment of the memory clinic visit, they discounted IDK tokens appearing in relatively standardized question sequences such as the Mini-Mental State Examination [MMSE [18]] which is commonly used by primary care providers [19]. Analysis of precisely these IDK instances could be especially informative for practitioners. By providing principled criteria for consistent IDK interpretation, a function-based analysis of IDK use in formal cognitive assessments has the potential to improve reportedly questionable [20] test reliability.

In order to further assess the diagnostic value of IDK in the context of **cognitive impairment** consultations, we present a quantitative, function-sensitive analysis of patients' IDK production in MMSEs. In doing so, we provide objective, trainable patterns for IDK interpretation in **cognitive impairment** assessment. Building on Mikesell [11] and Elsey et al. [14], we hypothesize that proportionally increasing use of lack-of-knowledge (LOK) IDK (as opposed to its use as a non-cognitive interactional device, see Section 3.1.2) will correlate with higher **physician-assessed cognitive impairment** severity. Our quantitative evaluation and descriptive elaboration of this relationship will add another tool to the **cognitive impairment** assessment repertoire, thus assisting clinicians in the task of recognizing and stratifying **cognitive impairment**.

2. Methods

2.1 Sample

We analyzed 72 audio-recorded and transcribed physician-patient interactions selected from a large database of naturally occurring ambulatory care visits: the Verilogue corpus [21]. **Our secondary data analysis of the selected Verilogue interactions was exempted from further review as non-human research by the Michigan State University Institutional Review Board (IRB# x12-362e/ APP# i040882). All US-recorded interactions meeting the following criteria were included in our sample:** 1) the physician submitting the recording identified 'dementia' as the primary condition being addressed during the visit; 2) the physician assigned a subjectively determined severity level ('mild', 'moderate', or 'severe') to the patient's **cognitive impairment**; 3) the visit included at least a partial MMSE administration as defined by the use of standardized assessment questions; and 4) the patient produced at least one transcribed instance of "I don't know" in response to an MMSE question (as determined by an automated search for the string "don't know" co-occurring in a patient-uttered turn with an overt or implicit "I" subject pronoun). The included interactions were recorded from 2009 to 2013, and submitted by both neurologists and primary care physicians. Patient demographic characteristics for our sample are given in Table 1. Exhaustive manual extraction of all patient-produced IDKs from the MMSEs in these interactions yielded a final set of 189 fully codable IDK tokens. **Though an *a priori* power analysis could not be performed for lack of reliable effect size estimates in the literature (i.e., there were no studies reporting distributions of IDK across cognitive impairment severity levels), this sample size is comparable to those reported in previous quantitative studies of IDK variability [10,22].**

[TABLE 1]

2.2 Coding

2.2.1 Linguistic coding

The two authors, both trained variationist linguists with prior experience in IDK analysis [10,12,22], coded each of the 189 IDK tokens for a series of linguistic factors: 1) phonetic form; 2) boundedness **(i.e. association with additional content in the same clause)**; 3) prosody **(i.e. stress patterns conveyed by emphatic use of volume and vowel quality)**; 4) pronoun presence/absence; 5) adverbial modification; and 6) discourse function. Codes for 1-5 are given along with illustrative examples in Table 2; discourse function is detailed in Section 3.1. **All 6 of these factors have been shown to affect the distribution of IDK in prior work [10].** The first author was responsible for coding factors 1-5. She performed iterative passes through the data focusing on one variable at a time whilst being blinded to her prior codings of other variables. In rare cases of unclear assignment, almost exclusively pertaining to phonetic form, the authors discussed and jointly assigned the token. The use of a single, primary coder can be justified for these variables as the same coders have achieved inter-rater reliabilities in the 87-98% range in prior IDK datasets [22]. In contrast to IDK's linguistic characteristics (1-5), IDK's discourse functionality is known to vary between conversational contexts [22]. Thus, to yield accurate descriptions of impaired patients' IDK usage in MMSEs, a *de-novo* function coding protocol was established by means of an in-depth qualitative analysis **(see Section 2.3).**

[TABLE 2]

2.2.2 Social coding

Following the **linguistic coding process**, each IDK token was associated with the following social factors provided as meta-data in the Verilogue corpus: patient gender (male vs. female); age category (≤74 vs. ≥75); and severity of **cognitive impairment** ('mild' vs. 'moderate' vs. 'severe'). Age and gender have been shown to affect IDK distribution in previous research [10]; severity of cognitive impairment served as our key independent variable in the quantitative analysis. **Patient age was given as ≤74 vs. ≥75 of age in the**

database for privacy reasons and could not be recoded into more granular categories. Individual MMSE scores were not provided as meta-data and could not be consistently coded from the interactions by the authors due to the use of modified or abbreviated versions of the test. In order to test our hypothesis that patients' IDK use changes with disease progression, we therefore had to rely on the severity assessments participating physicians provided for patients' cognitive impairment. These assessments were provided by physicians unaware that future analyses like ours might rely on them for research purposes.

2.3 Qualitative analysis

IDK's discourse functionality varies across contexts of use [22]. It would thus be inappropriate to apply existing functional taxonomies derived from analysis of IDK in contexts other than MMSEs to the present dataset. Instead, we undertake a data-driven qualitative analysis to identify the functions and contextual patterns associated with cognitively impaired patients' use of IDK in cognitive assessments. The purpose of this exercise is two-fold: 1) to allow for the incorporation of discourse function into our quantitative analysis; and 2) to provide researchers, health providers and caregivers with a detailed taxonomy of IDK use in the specific context of MMSEs.

Our approach to qualitatively analysing IDK function in our data is informed by Conversation Analysis [23]. Multiple factors combine to determine the functionality of IDK in interaction: discourse context, sequential position, turn position, prosody, intonation **(i.e. relative pitch)** and paralinguistic phenomena. In line with the conversation analysis literature, we prioritise the role of sequential and turn position in isolating those IDK tokens that do not signal lack of knowledge, i.e. those acting as interactional devices to signal speaker stance or affect discourse progression (henceforth non-LOK IDK). We thus interpret IDK function based on patient response structure, information relations between utterances [24], and other objectively identifiable turn elements. This contrasts with the heavy reliance on more subjective factors such as tone of voice, accompanying laughter) which feature strongly **in the qualitative analysis of IDK functionality** in other situational contexts [10,25].

For this study, the **initial qualitative taxonomy** was developed by the second author, whose doctoral research had focused specifically on IDK function elucidation. The schema was subsequently tested for applicability and consistency by the first author, a doctoral-level linguist and physician. This yielded an inter-rater reliability of 93% and affirmation that the schema was plausibly practicable for clinical use.

We report the results of our qualitative analysis in Section 3.1 through the discussion of representative examples and characterization of their contextual patterns. Though the primary motivation for example selection was to represent clearly identifiable themes in the data, preference was given to brief, self-contained excerpts for considerations of space. A concise summary of our functional taxonomy is given in Table 3.

2.4 Quantitative analysis

In order to test our hypothesis that **cognitive impairment** severity significantly predicts MMSE IDK function in our dataset **and assess the contribution to IDK function of other relevant contextual factors, the following predictors were incorporated into a multiple logistic regression in the R Statistical Environment [26]:** phonetic form **(full vs. reduced)**, syntactic boundedness **(bound vs. unbound)**, prosody **(stressed vs. unstressed)**, patient age **(<75 years of age vs. ≥75 years of age)**, patient gender **(male vs. female), and severity of patients' cognitive impairment (mild vs. moderate vs. severe)**. Multiple logistic regression modelling is a preferred technique for the evaluation of correlation-based hypotheses in quantitative variationist analysis [27,28], given that the data representing all potential predictors of the dependent variable (in this case MMSE IDK function) are well-distributed with respect to one another. **To ensure adequate cell sizes for quantitative analysis and identify potential relationships**

between our coded predictors, we performed cross-tabulations of all linguistic, social, and function-based codes. This exploratory data analysis demonstrated that pronoun use and adverbial modification were too poorly distributed to be included in the model, owing to the infrequency of pronoun absence and modification presence. **All other factors had sufficient cell counts and were modelled using treatment or sum contrasts as appropriate. Sum contrasts were used to model severity based on the ordinal nature of the variable. These potential predictors were tested for their effect on our dependent variable, IDK functionality, dichotomized as LOK (see Section 3.1.1) and non-LOK (see Section 3.1.2).**

3. Results

3.1 Qualitative results

This section introduces the position-based taxonomy for isolating LOK from non-LOK uses of IDK. It differentiates seven sub-types of IDK use in MMSEs which are illustrated with examples reproduced *verbatim* from the Verilogue corpus (see the Appendix for a key to the transcription conventions).

3.1.1 LOK uses of IDK

Cognitively impaired patients' knowledge disavowal uses of IDK fall into four broad categories, identifiable by their differential positions within **patient** responses and in relation to **physician** questions as well as their subtly differentiated interactional effects. In (1), the **physician** asks the **patient** for the name of his clinic. The **patient** immediately follows the **physician**'s question with IDK to signal her knowledge gap and inability to provide the name of the clinic. A **patient** knowledge-gap is also communicated by IDK in (2). Here, however, IDK is preceded by an unfilled pause, suggesting that the **patient** makes an effort to recall the requested information before declining her knowledge of it. As illustrated, LOK IDKs in the position immediately following **physician**s' MMSE questions occur without (1) and with dependent complements (2). We classify both of these LOK IDKs as **isolated response** in our taxonomy, as the IDK-related content (when provided as in (2)) is syntactically dependent on IDK. Unlike our other subclassifications, in (1) and (2), the **patient**'s preceding and subsequent utterances do not serve to substantiate or otherwise modify the **patient**'s knowledge disavowal.

(1) DR: Okay. And what's the name of this clinic, more or less.
PT: ***I don't know***.
(Verilogue, 29256, 105-109)

(2) DR: Alright. So what's today's date.
(..)
PT: ***I don't know*** today's date.
(Verilogue, 33388, 56-58)

Example (3) illustrates a different type of IDK-related content: a discourse-coherent string. In this case the **patient**'s turn-initial IDK is associated with a syntactically independent memory account, i.e., a face-saving strategy used by impaired patients to 'explain, justify or excuse memory performance' [6]. IDK signals a knowledge gap, and the following turn elements provide a justification for it: the **patient** does not, as a habit, consult her calendar which is why she is unable to provide the date. The **patient** thereby implies an effect-and-cause discourse coherence relation between IDK and the string *I never look at the calendar* [24]. The memory account thus reinforces the knowledge disavowal meaning of IDK, a relationship that we label as an **account response** in our LOK IDK taxonomy.

(3) DR: First of all, what's the date today?
PT: ***I don't know***. (.) <I really don't. I never> look <@ at the calendar.@> @
(Verilogue, 50173, 143-144)

Another syntactically independent, discourse position-based indicator of IDK's knowledge disavowal meaning is the positioning of IDK after previous **patient** declarations of insufficient knowledge. Similarity discourse-coherence relations [24] relate IDK to such declarations in **patient**s' multi-unit MMSE response sequences. For example, the **physician** in (4) asks the **patient** to name the three words he had earlier asked her to remember. While the **physician** encourages the **patient** to activate her memory in an attempt to recall the three words, the **patient** repeatedly declares her inability to do so (*My brain is going to hell. I can't remember now! I can't remember everything!*). The two knowledge-disavowal tokens of IDK at the end of extract (4) reinforce the **patient**'s previously declared inability to provide a list of the three words. We label these lack of knowledge-reinforcing tokens **final response** LOK IDKs.

(4) DR: Okay. Name those three things.
PT: *(.h)* Oh geez. Oh geez. >May brain is going to hell.<
DR: == Take your time. Take your time.
PT: *(sulky voice)* I can't remember now!
DR: Okay. I pointed to three things. There are two things, and name one other thing.
PT: *(sulky, forceful voice)* I can't remember [everything!]
DR: [Okay.] Alright. Okay.
PT: **I dunno**.
DR: == Okay.
PT: **I [don't know**.]
DR: [That's okay.] That's okay. Alright.
(Verilogue, 50357, 129-139)

Lastly, turn-final IDK functions to signal insufficient knowledge when it follows either a partially correct or aborted patient answer. Here, IDK is related to preceding answer attempts in a violated expectation discourse-coherence relation [24]: patients give the impression that they will be able to give a complete answer to the MMSE question but use IDK to cancel this assumption. In (5), the patient provides two of the three words she had been asked to recall (*Chicago*, *Cadillac*). Following the continuation signal *and* [29], she says *I dunno* to signal the end of her turn and her inability to provide the third word. In (6), the patient uses IDK after providing a partially correct answer (or guess) (*20 something*) to signal that she is unable to provide a more precise answer than the one already given. In addition to signalling lack of knowledge, these tokens signal patients' desire to terminate the ongoing question sequence. Hence we describe them as **terminating response** LOK IDKs.

(5) DR: So what were the three words? (..) Remember the three words? What were they?
(.)
PT: Oh, Chicago and (.) Cadillac and, ***I dunno***.
(Verilogue, 37107, 91-92)

(6) DR: And what year is it?
PT: 190 (...) No, it's 20 something. ***I don't know***.
(Verilogue, 33388, 58-59)

Table 3 summarises our taxonomy of IDK LOK uses in MMSEs. As shown, LOK IDK occurs in distinctive combinations of positions and discourse-coherence relations to convey subtly different interactional effects.

[TABLE 3]

3.1.2 Non-LOK uses of IDK

As is the case with IDK use in other situational contexts (see Section 1), IDK is regularly used in MMSEs for non-cognitive functions. We divide these non-LOK IDK uses into three categories based on their sequential and turn position as well as their interactional effects.

Unlike the tokens in (1)-(6), IDK in (7) does not signal LOK but serves to qualify the patient's commitment to her MMSE response. Contrast discourse-coherence relations [24] relate the non-LOK IDK token to the patient response, acting as a non-verbalized *but* (e.g., "Oh, I don't know. [but] I guess we're still in {city}.") regardless of the correctness of the patient's answer. Because they serve to hedge patients' MMSE responses, IDK tokens of this kind are termed **hedging use** non-LOK IDK in our taxonomy. They either precede or follow answers to MMSE questions.

(7) DR: Okay. What city are we in?

PT: Oh, ***I don't know***. I guess we're still in {city}.

(Verilogue 48500, 92-93)

When IDK tokens occur turn-medially in patient MMSE responses, as in (8), they function to link two parts of a turn that are otherwise unrelated. Crucially, it is the strings surrounding IDK that signal the patient's inability to supply the requested information, not IDK itself. IDK conveys the patient's communicative presence while she is working to recall the information requested by the **physician**. Based on IDK's role as a bridge between otherwise unrelated turn elements, we label these tokens **bridging use** non-LOK IDKs.

(8) NR: Do you know what day of the week?

(…)

DR: You know, Monday, Tuesday, Wednesday, Thursday.

PT: <@Yes, I know @ I know what that. @> @ ***I don’t know*** I just go by (.) by what is today? Friday?
NR: Today’s Tuesday.
PT: Today’s Tuesday. You see. ***I dunno***. I just go by today. Nobody ever asks me what day it is.
(Verilogue 28996, 22-28)

Finally, non-LOK IDKs function to signal patients’ *resistance*, rather than *inability*, to answer **physician**s’ MMSE questions. In (9), the patient does not use IDK to signal her inability to list animals but her reluctance to engage with the task initiated by the **physician** question. Though *which ones to start with* represents syntactically related IDK content, there is no meaningful discourse coherence relation (only perhaps a semantically bleached *and*) linking IDK to its preceding string *there’s so many*. As with the bridging use non-LOK IDK, this IDK use is distinguishable in part by its lack of a characteristic discourse-coherence relation with the surrounding patient talk. It redirects the conversation away from the resisted MMSE question and is thus called **redirecting use** non-LOK IDK in our taxonomy.

(9) DR: I want you to name as many animals as you can in a minute. […] You can start.
(.)
PT: Oh. (.) u:m (..) there’s so many <@ different. ***I dunno*** @> which ones to start with.
(Verilogue, 29940, 88-89)

Table 4 below summarizes our taxonomy of non-LOK uses of IDK in MMSEs. It illustrates how these IDK uses are uniquely identifiable by the combination of their characteristic discourse positions and either the presence or absence of discourse-coherence relations with surrounding patient talk.

[TABLE 4]

3.2 Quantitative results

The majority (83%, N=156) of our MMSE IDK tokens were identified as performing a LOK function (**Isolated response**, **Account response**, **Final response**, or **Terminating response**). Our multiple logistic regression analysis predicting LOK (as opposed to non-LOK) IDK functions yielded a single significant predictor: severity (see Table 5). **'Severe' cognitive impairment was positively associated with LOK IDK use relative to 'mild' cognitive impairment (p=0.01). The odds ratio for LOK IDK use given 'severe' cognitive impairment was 2.24, 95% CI=(0.18, 1.44). 'Moderate' cognitive impairment did not significantly predict LOK IDK production (vs. 'mild', p=0.85, OR=0.95, 95% CI=[0.54,1.69]). Likewise, phonetic form (p=0.29), syntactic boundedness (p=0.60), prosody (p=0.92), patient age (p=0.83), and patient gender (p=0.26) did not demonstrate significant differences in their distribution between LOK vs. non-LOK coded tokens. Estimates, standard errors, odds ratios and confidence intervals are given for all factors in Table 5.**

Focusing on severity, MMSE IDK function is notably varied among 'mildly' **impaired** patients in particular. As shown in Table 5, only 68% (N=23/34) of the MMSE IDK tokens from 'mild' patients were LOK IDK, compared to 92% (N=46/50) of the MMSE IDK tokens produced by 'severe' patients. **Based on our qualitative analysis, the remaining 32% of the 'mild' patients' MMSE IDKs performed non-LOK functions. If all of the MMSE IDK responses considered in this analysis were uncritically interpreted as LOK IDK (i.e., as knowledge disavowals) and thus scored as incorrect or non-answers during MMSE administrations, 32% (N=11) of IDK responses by the mildly cognitively impaired would be mis-scored. Based on our sample, insensitivity to the LOK/non-LOK distinction would be proportionately less problematic for 'moderate' and 'severely' cognitively impaired persons, as only 17% of the tokens produced by the former and 8% of the tokens produced by the latter were identified as non-LOK. However, these groups typically generated more IDKs per MMSE, with means of 2.7 IDKs/patient for those labelled 'moderate' and 2.8 IDKs/patient for those labelled**

'severe', compared to only 1.9 for those labelled 'mild' (see Table 5 for LOK and non-LOK proportions across all tested factors).

[TABLE 5]

4. Discussion and conclusion

4.1 Discussion

The current study provides a mixed-methods analysis of MMSE IDK in **cognitive impairment** interactions, offering a position-based functional taxonomy of IDK uses specific to the setting of cognitive assessment as well as a quantitative sociolinguistic account of the distribution of macro-level IDK functions (LOK vs. non-LOK). This work enhances our understanding of assessment in **cognitive impairment** by demonstrating a relationship between objectively recognizable IDK features and **impairment** severity. Providers may utilize the qualitative and quantitative evaluation of IDK use provided here to increase their awareness of MMSE IDK variation and potentially improve the accuracy of their cognitive assessments, especially in 'mild' **cognitive impairment**.

Cognitive assessment is a socially complex, interpersonally difficult clinical task for which relatively few primary care providers profess confidence [4,9]. Focusing on a discrete, frequent, and diagnostically significant patient-produced phrasal feature [14]—MMSE IDK—we present a teachable means by which to attenuate the explicit face-threats occurring in memory tests. This has the potential to promote provider-patient rapport as well as provider confidence in the cognitive assessment process. For example, correct interpretation of patients' LOK IDK responses facilitates the coding of an incorrect or non-answer to an MMSE question; confident interpretation of these tokens avoids additional **physician** question reiterations or follow-up probes that may only serve to frustrate or embarrass **cognitively impaired** patients. Accurate identification of non-LOK IDK uses is equally necessary for productive, relationship-building dialogue. Patients use non-LOK IDK to buy time for additional consideration of the

MMSE question (**bridging use**) or avoid a difficult-to-answer MMSE question (**redirecting use**). Physicians restating the prompt or simply pausing may allow patients to produce a correct answer in these scenarios, thus reinforcing their positive self-identity and giving a more accurate impression of their cognitive functioning.

Moreover, our quantitative analysis underlines the clinical importance of distinguishing MMSE IDK uses based on function. The gradated difference observed in the proportion of LOK vs. non-LOK MMSE IDK in **cognitive impairment** severity groupings supports previous claims that IDK may be diagnostically useful in disambiguating subtle manifestations of **cognitive impairment** from functional memory disorders, mood disturbances, etc. Critically, our work provides a necessary nuance to Elsey et al.'s (2015) observation that incomplete and IDK-containing patient responses are characteristic of true **cognitive impairment**. We show that it is only the LOK category of MMSE IDK that associates with increasing severity. In fact, our results indicate that between **8% and 32% of MMSE responses (depending on the severity of the patient's cognitive impairment)** could be mis-scored if the rater was not sensitive to the multifunctionality of IDK and indiscriminately interpreted all instances of IDK as knowledge disavowals, i.e., verbal manifestations of cognitive malfunctioning. This is consistent with prior criticisms of the MMSE's lack of precision in real-world conditions [20]. Such imprecision could be clinically significant as the indications for some dementia medications are severity-dependent.

In the MMSE setting, severity is the only significant variable in predicting LOK IDK. This is in stark contrast to other medical and non-medical contexts where phonetic form, speaker age, and other sociolinguistic factors played a role in explaining functional variation [10,12,22]. What is particularly striking is the lack of a relationship between phonetic form and discourse function. In previously examined, non-impaired populations, LOK uses of IDK strongly correlated with phonetically full forms and non-LOK uses with phonetically reduced forms. Additional research is needed in both healthy ageing and pathologically ageing older populations to establish whether the absence of otherwise robust IDK form-function correlations in our MMSE data is indicative of cognitive deterioration, as suggested by our results, or of non-pathologic sociolinguistic change occurring over individuals' lifespans.

4.1.1 Limitations

Though our study is the only quantitative study of MMSE IDK function in naturally occurring **cognitive impairment** consultations, it nonetheless relies on a relatively small convenience sample of patients. As a result, it could be underpowered to detect correlations in IDK use. Importantly, however, our sample is comparable to or larger than those used in previous IDK work, as described by Pichler [10]. **We did not have access to quantitative MMSE scores or exact patient ages. An ideal sample for this analysis would include these data and follow patients longitudinally, tracking their MMSE IDK production as their cognition deteriorates.** Our cross-sectional findings support future research utilizing multiple data collection time-points; our study design and execution provide a replicable framework for conducting such research.

4.2 Conclusion

This combined conversation analytic and quantitative variationist analysis of MMSE IDK production by cognitively impaired individuals in physician-patient visits identifies potentially trainable patterns in IDK use that correlate with physician-assessed dementia severity. It illustrates objectively distinguishable IDK functions derived through a context-specific, data-driven process. Furthermore, it demonstrates that lack of knowledge (LOK) IDK uses (as opposed to non-LOK IDK uses) are associated with more severe clinical manifestations of dementia, and that mildly cognitively impaired individuals in particular produce a characteristically varied mix of IDK functions in their MMSE exams comparable to that of non-impaired adults [22].

4.3 Practice implications

Informed recognition of IDK's macro- and micro-functions encourages neutral or even rapport-building dialogue in the context of explicit memory testing. Providers may use IDK function variability, defined objectively in our position-based MMSE-specific functional schema, as a supportive indicator of dementia severity during these assessments, thus potentially improving their diagnostic accuracy in clinically ambiguous scenarios.

APPENDIX

Key to transcription conventions

DR	physician speech
PT	patient speech
[ ]	overlap
==	latching
=	turn continuation
(h)	inbreath
@	laughter
> <	reduced tempo
CAPITALS	louder than surrounding talk
underlining	emphatic stress
:, ::, :::, ::::	short, medium, long and very long syllable lengthening
(.), (..), (…)	short, medium and long pause
.	final intonation contour
,	continuing intonation contour
?	rising intonation contour
(text)	uncertain transcription
(?)	undecipherable words
[text]	extra-linguistic information
[…]	text omitted
bold italics	used in examples to highlight the feature discussed in the text
{text}	de-identified information

ACKNOWLEDGEMENTS

The authors contributed equally to the production of this manuscript and the analyses reported here. We did not receive any assistance in writing or data analysis. We gratefully acknowledge the generous financial assistance of the *Newcastle University Faculty for Humanities and Social Sciences*, *Spectrum Health* and the *Michigan State University Graduate School*. We appreciate the insightful comments and questions from International Conference on Communication in Healthcare 2015 attendees, from our anonymous reviewers and the editor. Finally, we would like to thank Virginia Stever for her assistance in testing our coding scheme.

REFERENCES

1. C. Vassilas, J. Donaldson, Telling the truth: what do general practitioners say to patients with dementia or terminal cancer? Br J Gen Pract 48 (1998) 1081–2.

2. H. van Hout, M. Vernooij-Dassena, K. Bakkerb, M. Blomc, R. Grol, General practitioners on dementia: tasks, practices and obstacles, Patient Educ Couns 39 (2000) 219–225.

3. O. Karnieli-Miller, P. Werner, J. Aharon-Peretz, S. Eidelman, Dilemmas in the (un)veiling of the diagnosis of Alzheimer's disease: Walking an ethical and professional tight rope, Patient Educ Couns 67 (2007) 307–314.

4. R.T. Woods, E. Moniz-Cook, S. Iliffe, P. Campion, M. Vernooij-Dassen, O. Zanetti, Dementia: issues in early recognition and intervention in primary care, J R Soc Med 96 (2003) 320–324.

5. P. C. Kontos, Embodied selfhood: An ethnographic exploration of Alzheimer's disease, in: A. Leibing, & L. Cohen (Eds.), Thinking about dementia: Culture, loss, and the anthropology of senility, Rutgers University Press, New Brunswick, 2006, pp. 195–217.

6. P.A. Saunders, K. de Medeiros, A. Bartell, 'Oh he was forgettable': Construction of self identity through use of communicative coping behaviors in the discourse of persons with cognitive impairment, Dementia 10 (2011) 341–359.

7. J. Phillips, C. Dimity Pond, N.E. Paterson, C. Howell, A. Shell, N.P. Stocks, S.M. Goode, J.E. Marley, Difficulties in disclosing the diagnosis of dementia: a qualitative study in general practice, Br J Gen Pract 62 (2012) e546-e553.

8. E. Tullo, L. Allan, What should we be teaching medical students about dementia? Int Psychogeriatr 23 (2011) 1044-1050.

9. D.J.F. Mayne, L. Allan, E. Reynish, A.M.J. MacLullich, E.R.L. Cunningham Vardy, Experience and opinions on post-graduate dementia training in the UK: a survey of selected consultant geriatricians, Age Ageing 43 (2013): 263–266.

10. H. Pichler, The Structure of Discourse-Pragmatic Variation, vol. 13, John Benjamins Publishing, Amsterdam, 2013.

11. L. Mikesell, Conversational practices of a frontotemporal dementia patient and his interlocutors, Res Lang Soc Interact 42 (2009) 135–62.

12. A. Hesson, H. Pichler, Breaking down barriers: interpreting patients' use of I DON'T KNOW in pediatric mental health consultations, M.S.

13. I. Hutchby, Resisting the incitement to talk in child counselling: aspects of the utterance ‘I don’t know’, Discourse Stud 4 (2002) 147–168.

14. C. Elsey, P. Drew, D. Jones, D. Blackburn, S. Wakefield, K. Harkness, A. Venneri, M. Reuber, Towards diagnostic conversational profiles of patients presenting with dementia or functional memory disorders to memory clinics, Patient Educ Couns 98 (2015) 1071–1077.

15. W. A. Beach, T.R. Metzger, Claiming insufficient knowledge, Hum Commun Res 23 (1997) 562–588.

16. J. Potter, Discourse analysis as a way of analysing naturally occurring talk, in: D. Silverman (Ed.), Qualitative Research. Theory, Method and Practice, Sage, London, 2004, pp. 200–221.

17. A.B. Tsui, The pragmatic functions of *I don't know*, Text 11 (1991) 607–622.

18. M.F. Folstein, S.E. Folstein, P.R. McHugh, Mini-mental state: a practical method for grading the cognitive state of patients for the clinician, J Psychiatr Res 12 (1975) 189–98.

19. R. J. Davey, S. Jamieson, The validity of using the mini mental state examination in NICE dementia guidelines, J Neurol Neurosurg Psychiatry 75 (2004) 343-344.

20. P. Bowie, T. Branton, J. Holmes, Should the Mini Mental State Examination be used to monitor dementia treatments? Lancet 354 (1999) 1527-1528.

21. J. Kozloff, J. Barnett, Verilogue corpus of patient-provider interaction, Verilogue, Inc., 2006, www.verilogue.com.

22. H. Pichler, A. Hesson, Discourse-pragmatic variation across situations, varieties, ages: I DON'T KNOW in sociolinguistic and medical interviews, Language and Communication 49 (2016) 1–18.

23. J.M. Atkinson, J. Heritage, Structures of Social Action. Studies in Conversation Analysis, Cambridge University Press, Cambridge, 1984.

24. F. Wolf, F.E. Gibson, Representing discourse coherence: A corpus-based study, Computational Linguistics, 31 (2005) 249-287.

25. H. Pichler, The functional and social reality of discourse variables in a northern English dialect: I don’t know and I don’t think compared, Intercultural Pragmatics 6 (2009) 561-596.

26. R development core team, R: A language and environment for statistical computing, Vienna: R Foundation for Statistical Computing, 2015, http://www.R-project.org.

27. D.E. Johnson, Getting off the GoldVarb standard: introducing Rbrul for mixed-effects variable rule analysis, Lang Linguist Compass 3 (2009) 359–383.

28. S. Tagliamonte, Analysing sociolinguistic variation (Key Topics in Sociolinguistics). Cambridge University Press, Cambridge, 2006.

29. D. Schiffrin, Discourse markers, Cambridge, Cambridge University Press, 1987.

TABLES

Table 1. *Sample demographics at the subject level, total subject N=72.*

Parameter	Value	$N_{subjects}$	Proportion
Gender	Female	47	65.3
	Male	25	34.7
Race	Non-white	13	18.1
	White	59	81.9
Age	55-74 years	20	27.8
	75+ years	52	72.2
Severity level	Mild cognitive impairment	18	25.0
	Moderate cognitive impairment	39	54.2
	Severe cognitive impairment	15	20.8
Home circumstance	Lives alone	11	15.3
	Lives in extended care facility	9	12.5
	Lives with caregivers/ family	52	72.2

Table 2. *Linguistic factors coded for the final sample of N=189 codable IDK tokens. Factors above the heavy black line were included in regression modelling (see 3.2); factors below the heavy black line could not be modelled (see 2.3). Function codes are exemplified in Section 3.1 and Tables 3-4.*

Factor	Codes	Description	Example
Form	Full	Discrete morpho-phonemic boundary between *don't* and *know*	*I don't know*
	Reduced	Reduced vowel in *don't* and/or no audible morpho-phonemic boundary between *don't* and *know*	*I dunno*
Boundedness	Bound	Followed by a dependent WH-word or a phrasal/clausal complement	*I don't know why* *I don't know what that means*
	Unbound	Not followed by overt complementation	*I don't know Ø*
Prosody	Stressed	Prosodic stress on either *don't* or *know*	*I don't know* *I don't know*
	Unstressed	No prosodic stress on *don't* or *know;* sometimes stress on *I*	*I dunno*
Pronoun use	Present	Audible vowel representing *I* immediately preceding the *don't know* string	*I dunno*
	Absent	No audible vowel conceivably representing *I* immediately preceding the *don't know* string	*Dunno*

Modification	Modified	Adverbial modification with *just*, *really*, *even* inside the IDK string	*I just don't know* *I don't really know*
	Unmodified	No adverbial modification inside the IDK string	*I Ø don't know*

Table 3. *Position-based typology of LOK IDK uses in MMSEs [DR = clinician; PT = patient; IDK = I DON'T KNOW; *n =* ***variable number of repetitions****].*

function	position-based distribution	discourse-coherence relation	gloss	example in Section 3.1
Isolated response	DR question + (pause) + PT IDK	(none)	*I cannot produce the requested information.*	(1), (2)
Account response	DR question + (pause) + PT IDK + PT account	effect and cause	*How would I know that?*	(3)
Final response	DR question + (pause) + PT IDK (*n) + DR question + (pause) + PT IDK	similarity	*I don't know the answer and I mean it.*	(4)
Terminating response	DR question + (pause) + PT answer (*n) + (pause) + PT IDK	violated expectation	*Please move on.*	(5), (6)

Table 4. *Position-based typology of non-LOK IDK uses in MMSEs [DR = clinician; PT = patient; IDK = I DON'T KNOW]*

function	position-based distribution	discourse-coherence relation	gloss	example in Section 3.1.2
Hedging use	DR question + (in)correct PT answer + PT IDK OR DR question + PT IDK + (in)correct PT answer	contrast	*Don't hold me to this.*	(7)
Bridging use	DR question + (correct or incorrect) PT answer + PT IDK + (correct or incorrect) PT answer	(none)	*I'm buying time to think.*	(8)
Redirecting use	DR question + (PT comment +) PT IDK + PT redirect	(none)	*I don't see the relevance of answering this question/completing this task.*	(9)

Table 5. *Multiple logistic regression model predicting LOK IDK functions based on linguistic and social predictors [CI= 95% confidence interval; * =statistically significant (alpha of 0.05); LOK= lack of knowledge].*

	LOK (vs. non-LOK function)				
grand mean 'LOK'	0.83				
total N	189				
deviance	164.8				
	N_{LOK}/N_{total} (%LOK)	**Estimate**	**Standard error**	**Odds ratio**	**CI**
Phonetic form p=0.29					
Reduced	72/91 (79.1%)	0.44	0.42	1.55	(0.69, 3.56)
Full	84/98 (85.7%)	(reference)			
Prosody p=0.92					
Stressed	127/154 (82.4%)	0.06	0.60	1.07	(0.34, 3.69)
Unstressed	29/35 (82.8%)	(reference)			
Syntactic boundedness p=0.60					
Unbound	110/133 (82.7%)	-0.27	0.52	0.76	(0.28, 2.18)
Bound	46/56 (82.1%)	(reference)			
**Severity *** p=0.011					
Severe	46/50 (92.0%)	0.81	0.32	2.24	(1.20, 4.23)
Moderate	87/105 (82.9%)	-0.05	0.29	0.97	(0.54, 1.69)
Mild	23/34 (67.6%)	(reference)			
Gender p=0.26					
Male	53/61 (86.9%)	-0.56	0.50	0.57	(0.20, 1.47)
Female	103/128 (80.5%)	(reference)			
Age p=0.83					
75 and older	113/137 (82.5%)	-0.10	0.46	0.91	(0.37, 2.32)
74 and younger	43/52 (82.7%)	(reference)			