# Linguistic markers of subtle cognitive impairment in connected speech: A Systematic Review protocol

Amélie B Richard, Manon Lelandais, Fabrice Hirsch, Sophie Jacquin-Courtois, Karen T Reilly

# Linguistic markers of subtle cognitive impairment in connected speech: A Systematic Review protocol

**Amélie B. Richard** (amelie.richard@univ-montp3.fr)
Paul-Valery Montpellier 3 University, CNRS UMR5267; Praxiling, 34090 Montpellier, France; Université de Lyon, Université Lyon 1, INSERM U1028; CNRS UMR5292; Lyon Neuroscience Research Centre, Trajectories Team, F-69676 Lyon, France https://orcid.org/0000-0002-5892-6215

**Manon Lelandais** (manon.lelandais@u-paris.fr)
Université Paris Cité, CLILLAC-ARP URP3967, 75006 Paris, France https://orcid.org/0000-0003-4962-0334

**Fabrice Hirsch**
Paul-Valery Montpellier 3 University, CNRS UMR5267; Praxiling, 34090 Montpellier, France https://orcid.org/0000-0002-5646-4651

**Sophie Jacquin-Courtois**
Université de Lyon, Université Lyon 1, INSERM U1028; CNRS UMR5292; Lyon Neuroscience Research Centre, Trajectories Team, F-69676 Lyon, France; Physical Medicine and Rehabilitation Department, Henry-Gabrielle Hospital, Hospices Civils de Lyon, 69610 Pierre-Bénite, France

**Karen T. Reilly** (karen.reilly@inserm.fr)
Université de Lyon, Université Lyon 1, INSERM U1028; CNRS UMR5292; Lyon Neuroscience Research Centre, Trajectories Team, F-69676 Lyon, France https://orcid.org/0000-0002-2840-1610

**Method Article**

**Keywords:** linguistic markers, cognitive impairment, connected speech

**Posted Date:** July 17th, 2023

**DOI:** https://doi.org/10.21203/rs.3.pex-2276/v1

# Abstract

**Introduction:** Mild Neurocognitive Disorders (mNCD) are characterized by a subtle decline in cognitive performance. Discourse analysis is a research method that investigates the verbal production of patients and has been used to screen for subtle cognitive impairment. Studies using discourse analysis are, however, rarely replicated. One reason for this is likely to be the lack of consistency across studies regarding the selection and definition of relevant linguistic features. The goal of this review is to identify reliable linguistic markers for the detection of subtle cognitive impairment in connected speech, and generate hypotheses associating these markers with underlying cognitive functions in a given discourse context.

**Methods and analysis:** Following the PRISMA guidelines, we will conduct a systematic review including five databases: PubMed, ScienceDirect, Embase, Web of Science, Google Scholar. We will include articles targeting a population of patients with subtle cognitive impairment, such as Mild Cognitive Impairment. Eligible articles will include at least one connected speech task and an analysis of linguistic features in this task. The data will be classified according to study population, connected speech task, analysed linguistic features, and methods used to compare cases and controls.

**PROSPERO registration number:** CRD42023394729.


# Introduction

# Rationale

Mild Neurocognitive Disorders (mNCD) are disorders characterized by a subtle decline in cognitive performance (American Psychiatric Association, 2013). This designation covers a large range of underlying pathologies such as dementia (e.g., early-stage Alzheimer's Disease, Lewy Body Disorders), vascular pathologies (e.g., stroke, traumatic brain injury), or treatment-related pathologies (e.g., cancer-related cognitive impairment).

Although subtle, mNCD can be very disabling and substantially decrease an individual's quality of life. Individuals with mNCD complain about forgetfulness, high distractibility, trouble with multitasking, and difficulties with learning and language (Petersen, 2016). Despite this, mNCD is often undetected by current neuropsychological tests, which are designed for more severe or advanced disorders. The lack of specific tools can lead to underdiagnosis and lack of access to appropriate treatments. There is thus a need to develop new tools to detect subtle cognitive changes.

Analyses focusing on speech and discourse are increasingly present in health research. This is likely due to their non-invasive character, high sensitivity, and ecological validity (Bryant et al., 2017). Indeed, these methods have yielded satisfactory results in the detection of individuals with mNCD compared to healthy controls (Boschi et al., 2017). Alterations in language might provide markers of subtle cognitive change because of its interaction with other cognitive functions. This is particularly true for language in context,

i.e., connected speech, which is characterised by sequences of spoken words organised according to phonological, syntactic and pragmatic rules. There is, however, little consensus about which linguistic markers are most informative regarding subject cognitive changes.

# Objectives

The aim of this review is to identify reliable linguistic markers for the detection of subtle cognitive impairment in connected speech and associate these markers with underlying cognitive functions.

## Reagents

## Equipment

## Procedure

The current protocol has been registered in the PROSPERO database (ID number: CRD42023394729).

# Eligibility criteria

## Language and date

Only papers where the full text is available in French or English will be included. Since our review focuses on mNCD, which was introduced into the literature with publication of the DSM-5 (2013), searches will be limited to the last ten years (2012-2023).

## Population

We will select studies that include healthy control subjects and patients with mild neurocognitive disorders (i.e., mile cognitive impairment (MCI), early-stage Alzheimer's Disease, Primary Progressive Aphasia, Parkinson's Disease, Multiple Sclerosis, Subjective Cognitive Decline, Right Hemispheric Stroke, Cortico Basal Syndrome, Behavioural Fronto Temporal Dementia, Lewy Body Dementia, Minor Stroke, Mild Traumatic Brain Injury).

We will exclude studies targeting individuals with major motor speech impairment (e.g., apraxia of speech) and/or individuals younger than 18 or older than 75. Studies with fewer than 10 participants in the patient group, with participants who score below the cut-off for mild cognitive impairment in screening tests (i.e., 21/30 for the Mini-Mental State Examination (MMSE) or 22/30 for the Montreal Cognitive Assessment (MoCA)), or without any control group will be excluded.

## Language tasks

We will include studies that include recorded and transcribed connected speech tasks, such as picture-based narratives, storytelling, or conversations. Studies with reading and/or writing and/or verbal fluency tasks and/or repetition tasks and/or naming tasks that do not contain connected speech tasks will be excluded.

## Neuropsychological tests

All studies must include at least one standardised neuropsychological test for screening cognition. Other neuropsychological tests will be reported if they are standardized and reported normalized scores.

## Methods

In this section, we will scope the analysed linguistic features and their linguistic domain (e.g., lexicon, phonetics), as well as the statistical methods used to compare cases and controls or/and the machine learning models used to classify populations and their outcomes (e.g., p-values).

## Publication type

Studies published in peer-reviewed journals will be eligible for inclusion. Grey literature (e.g., theses, preprints) and proceedings will not be analysed.

## Study design

Randomized Controlled Trials and case-control studies will be included. Meta-analyses, reviews or systematic reviews, case studies, qualitative studies, and methodological articles will be excluded.

# Search strategy

We will search published studies using electronic bibliographical databases. We hand search the references in included studies and during the selection phase we will add any relevant articles identified by our search alert.

The search equations will be re-run just before the final analyses and any newly-available studies will be retrieved for inclusion.

## Database

We will search the following databases: PubMed, ScienceDirect, Embase, Web of Science, Google Scholar.

To gain time, we will directly import the results from the online databases to Zotero and verify the metadata (e.g., whether the authors' names, publication date etc are correctly reported).

## Keywords

We will develop a search strategy based on keywords related to our review question. We will identify main keywords for the case population, connected speech tasks and methods. We will add synonyms to the search terms to extend our results. We will also add exclusion criteria such as "intervention" to avoid unrelated studies. With the help of Boolean operators, our prototypical syntax will be "case population" AND "connected speech task" AND "methods" NOT "exclusion criteria" (e.g., "early-stage alzheimer's disease" AND "picture-based description" NOT "therapy").

## Selection process

We will follow the PRISMA statement (Moher et al., 2009) and report our process in a PRISMA flowchart using an adaptation of the screening and selection procedure found in Pati and Lorusso (2018) and Mateo (2020). Pati and Lorusso (2018) propose a step-by-step method to conduct a PRISMA systematic review with extensive details and examples. Matteo (2020) describes helpful, easy-to-use methodological tools implemented in Excel and Zotero.

We will conduct our article selection following a three-stage procedure:

1. **Pre-screening:** This will focus on publication language, publication type, duplicates, incorrect metadata, and will be conducted by one author (AR).

2. **Title and abstract screening:** Studies that do not focus on connected speech will be excluded, this incudes studies that use naming tasks or verbal fluency tasks. The following exclusion criteria will be applied:

· Studies involving individuals with major motor disorders (e.g., apraxia of speech)

· Studies focusing exclusively on reading and writing tasks. While reading may be treated as spoken language, a strict definition would not classify it as connected speech. Similarly, writing is sometimes mistakenly classified as connected speech.

· Studies treating unrelated topics (e.g., qualitative studies in sociology focusing on discourse analysis and case studies). We expect to find studies with unrelated topics because of inclusion of the keyword polysemy.

This stage will be conducted by two authors (AR & ML). We will conduct an inter-rater agreement (%) analysis and aim to reach 80%. In the case of disagreement for the remaining 20% of the data, the two authors will examine the title and abstract a second time, discuss their eligibility, and try to reach a consensus. If no consensus is found, a third author (KR) will be asked to decide if the study should be included.

3. **Full-text selection:** We will assess each article's eligibility using the following exclusion criteria:

· Population: No control group; age below 18 or above 75 (mean); fewer than 10 participants per group.

· Tasks: Studies focusing on reading and writing tasks that do not include at least one connected speech task.

· Analysis: Studies that do not explicitly analyse speech parameters.

· Methods and study design: Descriptive statistics only, meta-analyses; systematic reviews, case studies; qualitative studies; theoretical methodological articles.

This part of the process will be carried out by two authors (AR & ML). We will conduct an inter-rater agreement (% and Cohen's κ) and aim for -near perfect agreement (κ > .81). Any remaining disagreement will be discussed with a third author (KR) until agreement is reached.

# Data collection process

## Data extraction and inclusion

Data from four categories will be manually extracted:

· Population: Number and age for both case and control populations.

· Tasks: Type involving connected speech.

· Tests: Neuropsychological tests used, cognitive domains tested.

· Methods: Linguistic features, statistical tests, outcomes.

In the case of missing information we will contact the authors to request the data. Unavailable data will be identified as such in our summary tables. All extracted data will be recorded in an excel database and summarised in narrative and/or table formats.

## Quality assessment

We will use the National Institute of Health Quality Assessment of Case-Control Studies scale to evaluate the quality of each study. Due to the nature of our data, we will also evaluate study quality on the basis of: the connected speech tasks, the linguistic variables, and the statistical tests used to compare the case and control populations. Two reviewers (AR & ML) will conduct a quality assessment giving each article a quality score (Poor, Fair, Good). We will compare scores using an inter- rater agreement (%) and discuss any remaining disagreement with a third author (KR).

# Synthesis methods

## Data synthesis

This section will provide information on the reliability of linguistic features as markers for subtle cognitive decline. We will incorporate the data in a narrative synthesis organised according to the study population, connected speech tasks, analysed linguistic features, and methods used to compare cases and controls. We will develop descriptive themes to describe the data, look for similarities and differences across studies, and when possible, merge similar data into higher-level categories and themes. For example, we will report the description of linguistic features and the measures for each article. We will then group and label those elements that share similar characteristics. We aim to provide easy-to-use guidelines to harmonise linguistic data across studies.

The reliability of a marker will be evaluated according to 1) the clarity with which it is described in the paper 2) Its ability to discriminate between individuals with mNCD and control subjects.

The data will be summarised in tables and figures depending on their category.

## Subgroup analysis

If appropriate, we will group studies according to the cognitive status of the participants, categorising patients based on scores in screening tests.

## Troubleshooting

## Time Taken

## Anticipated Results

The goal of this systematic review is to identify linguistic markers for subtle cognitive impairment in connected speech. We will present the main linguistic features found in the relevant literature, the methodology used to collect and process the language samples (i.e., connected speech tasks, statistical tests or machine learning methods), and the outcomes of the studies. Our aim is to use the results of this

review to make recommendations for the selection, classification, and reporting of linguistic features that are relevant for research in the field of subtle cognitive impairment. If possible, we will also provide guidelines for speech and language pathologists and therapists that can assist them in the often difficult task of identifying and assessing linguistic features in a clinical context. We expect our review to reveal gaps in the literature, and as such to provide avenues for future research.

## References


1. American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5e éd.).

2. Boschi, V., Catricalà, E., Consonni, M., Chesi, C., Moro, A., & Cappa, S. F. (2017). Connected Speech in Neurodegenerative Language Disorders : A Review. *Front. Psychol.*, *8*(269). https://doi.org/10.3389/fpsyg.2017.00269

3. Bryant, L., Spencer, E., & Ferguson, A. (2017). Clinical use of linguistic discourse analysis for the assessment of language in aphasia. *Aphasiology*, *31*(10), 1105-1126. https://doi.org/10.1080/02687038.2016.1239013

4. Mateo, S. (2020). Procédure pour conduire avec succès une revue de littérature selon la méthode PRISMA. *Kinésithérapie, la Revue*, *20*(226), 29-37. https://doi.org/10.1016/j.kine.2020.05.019

5. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses : The PRISMA Statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

6. Pati, D., & Lorusso, L. N. (2018). How to Write a Systematic Review of the Literature. *HERD: Health Environments Research & Design Journal*, *11*(1), 15-30. https://doi.org/10.1177/1937586717747384

7. Petersen, R. C. (2016). Mild Cognitive Impairment. *Continuum*, *22*(2), 404-418. https://doi.org/10.1212/CON.0000000000000313


## Acknowledgements

# Funding


This systematic review is funded by Université Claude Bernard Lyon 1, CNRS, INSERM, Lyon Neuroscience Research Centre (CRNL) U1028 UMR5292, TRAJECTOIRES Team, F-69500, Bron, France.

AR is financed by by a PhD grant awarded by Paul-Valery Montpellier 3 University and funded by the French Ministry of Higher Education (MESRI).