

A French Dataset for Multilingual Physical Commonsense Reasoning

Raphaël Merx
The University of Melbourne

1 Introduction

This paper describes the creation of a new French dataset for the MRL 2025 Shared Task at EMNLP on Multilingual Physical Commonsense Reasoning. The goal of this work is to provide a set of culturally and linguistically authentic examples that test a model’s understanding of physical properties, everyday common sense, and culturally specific knowledge within a French context.

The dataset was constructed by a single native speaker of French and follows the PIQA format, presenting a prompt and two possible solutions, one of which is correct. The contribution consists of over 100 unique, manually-created examples designed to be challenging and representative of real-world situations familiar to a native speaker.

2 Elicitation Process

The creation of the dataset followed an observational methodology to ensure a diverse and grounded set of examples. The elicitation process was driven by three main sources of inspiration:

- **Environment observation:** Many examples were derived from observing daily life and social interactions in various settings, such as at a café, at home, or in a park. This approach grounded the prompts in realistic, everyday scenarios.
- **Object-focused exploration:** A systematic exploration of a home environment yielded numerous prompts related to the function and physical properties of common objects. Particular attention was paid to tools found in the kitchen and to home improvement tools, leading to examples that require practical, hands-on reasoning.
- **External media inspiration:** French websites were browsed to generate ideas for more

general and culturally-specific topics. For example, the furniture and home goods website ‘leroymerlin.fr’ provided inspiration for objects/furniture, while the news outlet ‘leparisien.fr’ provided inspiration for current events, like sports.

3 Dataset content

The resulting dataset covers physical reasoning, general common sense, and cultural specificity.

Physical reasoning A significant portion of the dataset focuses on reasoning about the physical properties of objects and the outcomes of actions. These examples often require an intuitive understanding of physics. They include tasks such as how to properly use a drywall anchor, the most effective way to chill a bottle of wine, how to operate old shutters, and how to use the métro in Paris.

Common sense This category often overlaps with cultural norms and includes topics like traffic laws, basic cooking principles, and gardening tips.

Culturally specific knowledge I tried to include prompts that require knowledge specific to French culture, which might be opaque to outsiders. These examples cover:

- **Food:** Related to specific dishes, ingredients, or cooking tools.
- **Social norms:** Social etiquette when greeting, toasting glasses, and traditions associated with major holidays.
- **General culture:** Items related to popular sports like or iconic events.

Overall, the dataset mixes questions that are specific to French culture, with questions that are related to physical and commonsense reasoning, with some questions at the intersection of both (e.g. how to operate some objects commonly found in old French houses).