

To Eat and Beyond:

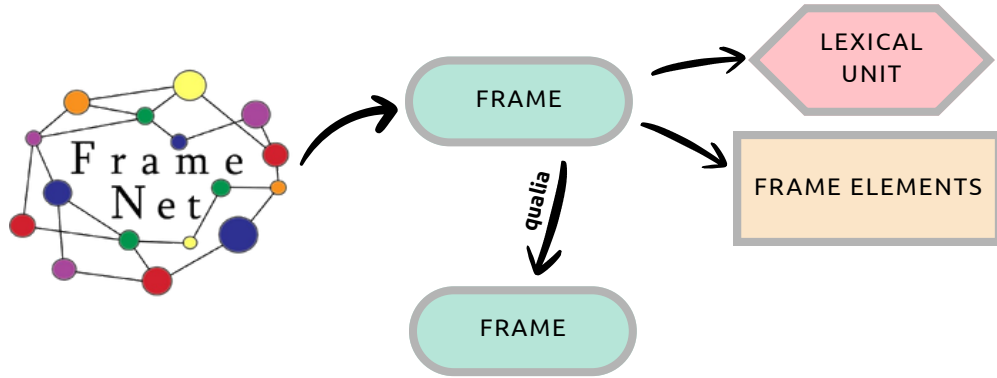
A FrameNet-Inspired Annotation of Food and Its Uses Over Time

Teresa Paccosi, Gauri Bhagwat, Marieke van Erp
KNAW Humanities Cluster DHLab, Amsterdam

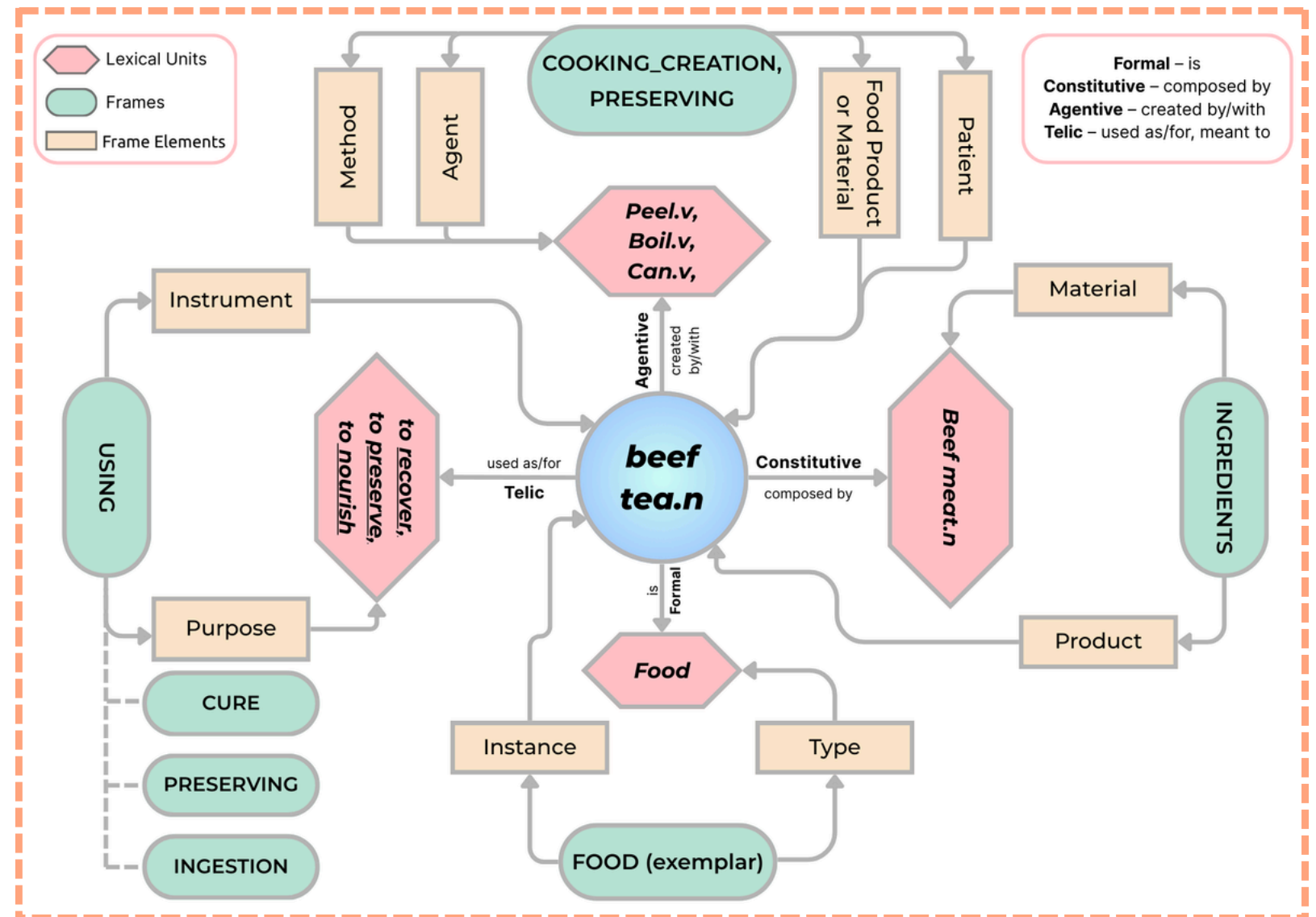
Motivation

Food plays a **central role** in everyday life, functioning not only as a basic human necessity but also as a rich domain where cultural norms, social practices, and personal identities are continually shaped. **Beyond mere nourishment**, examining both common and less frequent uses, such as medicinal or diplomatic practices, allows for a **more comprehensive understanding of the food domain**.

Capturing Food



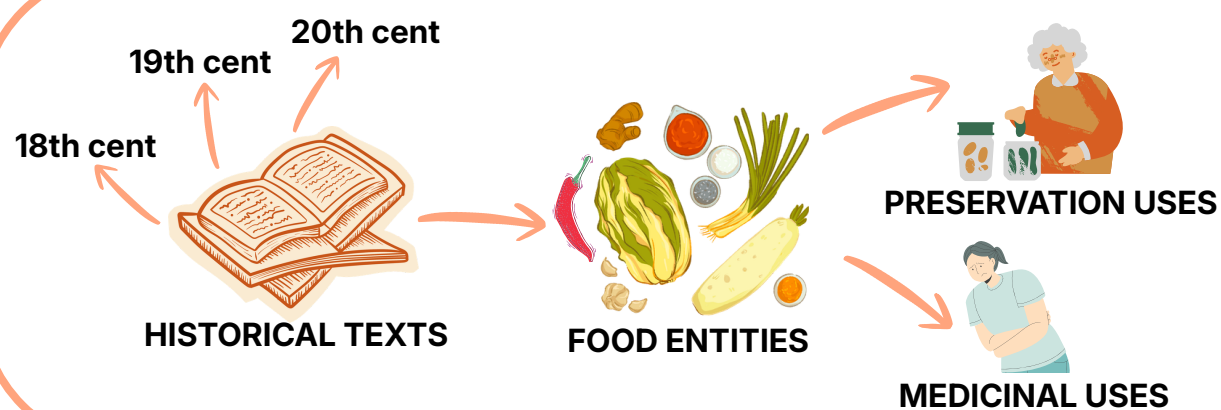
Our **scheme** captures food from multiple angles: composition, processes, and uses. LUs can serve as FEs in other frames, creating interconnected relations. This allows food to be represented more **comprehensively** as illustrated in the diagram.



Annotation Flow

- 1 **FOOD_LU** Ginger soothes sore throats and alleviates headaches.
- 2 **FOOD_LU** **CURE_LU** Ginger soothes sore throats and alleviates headaches.
- 3 **CURE_Food_Treatment** **FOOD_LU** **CURE_LU** **CURE_Affliction** **CURE_LU** **CURE_Affliction** Ginger soothes sore throats and alleviates headaches

Data Annotation



Semi-automatic process for document selection (**food entity extractor**+list of **LUs** for specific uses and contexts).

The final dataset includes texts from **three centuries** (18th, 19th, and 20th) and **three genres** (travel and ethnography, household and recipes, and medicine and science).

The dataset consists of approximately **6,000 annotated sentences** (~300,000 tokens).

Inter Annotator agreement (**IAA**) reached **0.94**, calculated on a sample of 100 randomly selected sentences.

Frame	Count
Cooking_Creation	444
Cure	545
Descriptor	1,240
Food	4,443
Ingestion	364
Ingredients	340
Preserving	576

Testing LLMs

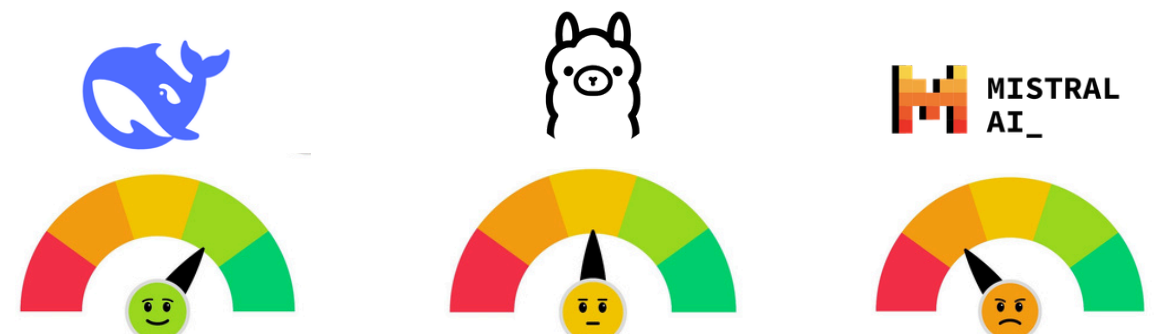
Frame	Deepseek-Distil-R1			Llama			Mistral			Deepseek_V3		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
PR	0.6182	0.6455	0.6205	0.2295	0.5791	0.3276	0.2164	0.3579	0.3204	0.7486	0.7459	0.7472
INGR	0.0220	0.0323	0.0262	0.0746	0.1667	0.1031	0.0495	0.1292	0.0641	0.2558	0.3667	0.3011
INGE	0.6254	0.6157	0.6214	0.4534	0.4310	0.4420	0.4505	0.3468	0.3114	0.7053	0.6667	0.6855
CURE	0.4635	0.4297	0.4589	0.3767	0.4861	0.4248	0.1786	0.4933	0.2518	0.5909	0.4432	0.5062
COOK	0.4221	0.3123	0.3515	0.1378	0.4598	0.2101	0.0711	0.3709	0.1192	0.1818	0.2292	0.2026
DESCR	0.3819	0.3433	0.3046	0.0933	0.2880	0.1426	0.027	0.3956	0.0425	0.2301	0.2886	0.2562
FOOD	0.5918	0.7543	0.6545	0.6148	0.7528	0.6778	0.1863	0.5487	0.2601	0.5864	0.7395	0.6540

Aim: Test clarity of guidelines and potential for scaling beyond manual annotation.

Test Set: 150–200 sentences per frame (excluding FOOD), balanced across genres and centuries.

Models: Reasoning-focused and instruction-tuned architectures of varying sizes.

Setup: Same prompt for all models with deterministic settings (temperature = 0.0, top-k = 1).



Conclusions

We introduced a Frame Semantics-based annotation scheme to model food and its historical uses, including less frequent practices such as medicinal and preservation-related applications. Experiments with LLMs and human agreement confirm both the **clarity of the guidelines** and the potential for **semi-automatic annotation**.

Future Work

Future work will expand annotations to other languages and genres, deepen collaboration with **food historians**, with the goal to develop a **benchmark** for automatically extracting food-related entities and uses in historical texts, enabling cross-linguistic quantitative analysis of diachronic corpora.

