

Chapter 7

Pragmatics as the Secret Ingredient: a multilingual study of reference tracking and evolving anaphors in culinary narratives

Anna Morbiato^{*, \diamond}

^{*}Ca'Foscari University of Venice

^{\diamond} Venice International University

Abstract

Anaphora resolution, namely the correct interpretation of anaphoric means, poses numerous challenges for natural language processing (NLP) technologies, particularly when approached from a multilingual perspective. Reference tracking mechanisms exhibit considerable cross-linguistic variation, with some languages heavily relying on inference and common sense. Additional complexities are introduced by phenomena such as null pronouns, partial coreference, and evolving anaphora, namely an expression that refers to an entity which, in the interim, has undergone changes in its properties or attributes. Crucially, these phenomena are frequent in procedural texts such as recipes, where precise coreference disambiguation is essential for the successful execution of tasks in a technologically driven culinary landscape. The present paper delves into the complexity and variation of reference tracking across languages NLP systems need to face. After an overview of coreference encoding and resolution from a cross-linguistic perspective, with a focus on zero, bridging, and evolving anaphors, it seeks to unravel the intricacies of anaphoric encoding in culinary narratives through a study of the type, frequency, and nature of anaphoric devices used in authentic recipe texts. It examines texts in English, Italian, and Chinese, which differ substantially in terms of coreference tracking systems and in the degree of reliance on inference for disambiguation.

Keywords: anaphora resolution, null pronouns, evolving anaphors, multi-language perspective, inference

7.1 Introduction

In the realm of natural language processing (NLP), recipes present a fascinating and challenging domain. Activities in the kitchen environment are particularly complex, usually involving many objects and interactions between them at different stages. Ingredients undergo rapid and substantial changes of state, getting moved and transformed throughout the process. Each step frequently leads to the creation of new items (resultant objects, see Morbiato and Cani, this volume), which are then subject to further movement and transformation, evolving over time as each subsequent step is carried out. Achieving a precise comprehension of recipe texts becomes pivotal in a technologically driven culinary landscape, especially with respect to when and how to use specific items during the various steps of preparation. Crucially, this comprehension comes with its set of challenges. Recipe texts abound with references to contextual items, including ingredients and kitchen tools, presupposing solid awareness and mastery of world knowledge and common sense. Moreover, they frequently incorporate multiple modalities, encompassing not only the text (the recipe itself) but also images (depicting ingredients and the cooking process) and, at times, audio, or video elements. Third, the linguistic means used in recipes to refer to items, ingredients, and tools are sometimes far from transparent and often include weak anaphoric means, such as pronouns or even null pronouns/zero anaphors (henceforth signalled as \emptyset). Ingredients and 'resultant objects' (i.e., items that are created as a result of a specific step, as e.g., a dough, or a puree, cf. Morbiato and Cani, this volume) are frequently left unmentioned, as exemplified in phrases like 'Bake \emptyset_x until \emptyset_x crispy and \emptyset_x golden' or 'Pour \emptyset_y until \emptyset_y saturated', resulting in potential ambiguity. This phenomenon is not marginal; in fact, zero anaphors (also referred to as 'zero pronouns', 'null elements' or, in the NLP literature, 'implicit entities') are frequent in procedural texts such as recipes, particularly when compared to other genres. Further challenges include partial coreference phenomena (also known as bridging) as well as so-called evolving or evolutive anaphors (the latter term is used by Asher (2000)), referring to the situation in which a textual reference is taken up by an anaphoric expression after undergoing radical transformations in its ontological state [36]. Evolving anaphors are common in culinary narratives, where they typically refer to a previously mentioned ingredient or component of the recipe, as happens in (1). Despite the fact that a change in properties occurs, as the pasta becomes soft after cooking, it is still referred to with a pronoun, *it*. Reference tracking might be even more challenging when a zero is used instead of an overt pronoun, as in (2). Crucially, inference plays a significant role in coreference disambiguation, in addition to, or in place of, lexical and grammatical encoding. It is world knowledge and common sense that enable the reader to understand that the *it* and the zero in (1) and (2) refer to the pasta, and not to the pot or the water.

- (1) Boil a pot of water and add the **pasta**. Once **it's** cooked, drain **it** and then add **it** directly to the tomato sauce.
- (2) Boil a pot of water and add the **pasta**. Once \emptyset cooked, drain \emptyset , and then add \emptyset directly to the tomato sauce.

Against this background, this paper focuses on the complexities of non-transparent anaphoric references in recipes, and it does so from a cross-linguistic perspective. After a cross-linguistic overview of coreference encoding and resolution, with a particular emphasis on zero pronouns, partial coreference phenomena, and evolving anaphors, it presents an analysis of the flow of information in recipe texts, as well as of the type, frequency, and nature of anaphoric devices used in authentic recipes drawn from food blogs. It considers texts in English, Italian, and Chinese, which significantly vary in terms of coreference tracking mechanisms and demonstrate different levels of dependence on zero anaphors and inferential processes.

The rest of the paper is structured as follows. Section 7.2 provides a cross-linguistic overview of reference encoding and tracking mechanisms, focusing on zero pronouns, partial coreference, and evolving anaphors, with few brief remarks on the challenges involved for NLP. Section 7.3 presents the analysis and its results. Lastly, Section 7.4 draws some conclusions.

7.2 Reference encoding and tracking

When talking about sequences of situations in which the same participants are involved, it is necessary to refer to them in each clause in such a way that they can be identified as being the same as or different from the participants referred to in previous clauses [21]. This helps granting textual cohesion, which concerns the way various elements within a piece of writing or speech are linked together to create a unified and understandable whole. The term ‘anaphora’ is derived from the Greek word *αναφορα*, a compound word consisting of *ανα* ‘back, upstream, back in an upward direction’, and *φορα* ‘the act of carrying’, denoting the act of carrying back upstream. It refers to a relation between two elements, wherein the interpretation of one (generally called ‘anaphor’) is determined by the interpretation of the other (generally called ‘antecedent’) [28]. In (3), the antecedent *onions* is then being referred to through weaker anaphoric means, i.e., the pronouns *them* and *they*.

- (3) Chop the **onions** finely. Then, sauté **them** in a pan until **they** turn golden brown.

Crucially, each language differs in the range of possibilities for creating anaphoric relations between elements in texts, which may be expressed through both morpho-syntax and lexicon [27]. Moreover, languages exhibit considerable variation in the extent to which grammar and inference play a role in coreference disambiguation. Awareness of such differences is paramount in NLP approaches that adopt a multi-language perspective. In some languages, disambiguation is facilitated through lexical information or syntactic structure, while in others, it heavily relies on inferential processes involving context, common sense, and real-world conceptual knowledge. Cross-linguistic variation observed in this domain has led scholars like Foley and Van Valin (1984) to outline a typology of reference tracking systems, which will be introduced in Section 7.2.1. Furthermore, the choice among different anaphoric encoding (noun repetition, substitution with hypernyms or synonyms, pronouns, or zeroes) is not random: the use of a pronoun over noun repetition may provide opposite instructions towards antecedent retrieval. This aspect will be illustrated in Section 7.2.2. Sections 7.2.3 and 7.2.4, finally, present a focus on anaphors that pose a particularly significant challenge for NLP—namely, zero, partial coreference, and evolving anaphors.

7.2.1 Reference tracking from a multi-language perspective

Typological approaches to language description typically identify four major types of reference-tracking systems, which a given language may use one or more of: (i) gender systems, (ii) switch-reference systems, (iii) switch-function systems, and (iv) inference systems [21, 14, 28]. The first type is commonly considered lexical, the second and third are grammatical, and the fourth is pragmatic. Italian, along with many other genetically different languages, is an example of the gender/class system: nominal phrases are gender/person/number marked, hence further mentions may be tracked through gender marking on pronouns or verbal affixes or by employing a combination of both strategies. While this phenomenon is observable only to a very limited extent in English, i.e., with personal pronouns (4), Italian’s gender and number

agreement morphology is significantly richer and extends to common nouns. In (5), the clitic masculine pronoun *-li* in the second clause involves coreference with the masculine noun *gambi* 'stalks', introduced by an oblique phrase, rather than to the feminine noun *fiori* 'flowers', thus overriding the semantic and syntactic hierarchies (i.e., object over oblique etc.). Crucially, two major translation services online (DeeepL, GoogleTranslate) provide the English translation in (a) which is, at best, ambiguous, in fact, incorrect: the pronoun *them* is interpreted as coreferent to the direct object 'the leaves', and not to 'the stalks'. To avoid this, a deictic like *the latter*, or a full NP repetition, is needed, as in the translation provided by ChatGPT in (b).

(4) When John_j met Clara_k again, she_k was willing to marry him_j.

(5) (source: recipe by internazionale.it/notizie/rachel-roddy/2018/12/21/consigli-alternativi-per-il-pranzo-di-natale, last accessed: 17/04/2024)

Pulite per bene le bietole [...],

'Clean thoroughly the chards [...],

poi separ-ate le fogli-e_j da-i gamb-i_k
then divide-IMP.2PL the.F.PL leaf.F-PL from-the.M.PL stalk-M.PL

e riduc-ete-li_k in due-tre pezz-i.
and cut-IMP.2PL-3PL.M in two-three piece-PL

a. # '[...] then separate the leaves_j from the stalks_k and cut them_j into two to three pieces. (Translation by DeeepL on 25/04/2024; a similar translation was provided by GoogleTranslate)

b. ' [...], then separate the leaves_j from the stems_k and cut the stems_k into two or three pieces.' (Translation by ChatGPT on 25/04/2024)

In switch-reference systems, found in languages of North America, Oceania, North Asia, and Africa [28], the verb of a dependent clause is morphologically marked to indicate whether the syntactic subject of that clause is the same as that of the main clause. Similarly, switch function systems track the reference of an NP by means of verbal morphology, which indicates the semantic function of that NP across clauses [21]. Cross-linguistically, the most common case is the active/passive alternation: in (6), the passive voice indicates that in the second clause the zero encodes the patient [28]:

(6) John_j killed Jack and Ø; **was** later **arrested** by the police.

In inference systems, finally, coreference disambiguation does not rely on lexical or grammatical cues, but rather on inference, pragmatics, and world knowledge. Such systems are characterised by substantial use of zero anaphora and are prevalent in East and South-East Asian languages such as Chinese, Japanese, Korean, Javanese, Tamil, and Thai [21, 28].

(7) C: 那个人把西瓜掉在地上碎了。

nà ge rén_j ba xīguā_k diào zài dì-shang Ø_k suì le
that CLF person PM watermelon drop at ground-on break PFV

E. That man dropped the watermelon on the ground and **it** burst.'

I. *L' uomo f-ece cadere l' anguri-a per terra e*
the.SG man.M.SG make-PST.3SG drop the.SG watermelon.F-SG for ground and
quest-a si r-uppe.
this-F.SG SELF break-PST

Sentence (7), also discussed in [3], illustrates the differences between English, Italian, and Chinese in coreference behaviour with conjunction reduction. English and Italian require overt

anaphoric means (the pronouns *it* and *questa*) which, thanks to gender agreement, signal that the reference is to be tracked back to the object of the first sentence (watermelon) and not the subject. In Chinese, no pronoun or overt anaphora is present. Inference and world knowledge suffice to get the hearer understand that the only possible item that bursts is the watermelon. In fact, the role played by inference, world knowledge, and context in reference tracking has been substantiated in several studies also for gender-system languages. Shared knowledge between the interlocutors or *common ground* has also been demonstrated to be crucial in anaphora disambiguation [31, 40, 18, 21, 22, 47, 11].

It has often been observed that variations in the encoding of anaphoric relationships pose challenges for NLP applications, such as Machine Translation: for example, when translating from and into languages which mark the gender of pronouns, resolving anaphoric relations becomes crucial [41], as exemplified in (5). According to Mitkov, complexity arises from gender or number discrepancies in words denoting the same concept, differences in the selection of anaphors in the target language, and ellipsis of pronouns in both the source and target languages.

7.2.2 Cognitive motivations underlying the choice of anaphoric means

It is widely accepted that the choice among different anaphoric devices is far from random and is in fact vital for allowing correct reference tracking [34, 35, 38]. Full noun phrases, for example, serve either to introduce brand new ('indefinite') referents into the discourse or to reintroduce old ('definite') referents after a significant absence, while weaker means (pronouns, zeroes) are used to encode informationally given/accessible items (example from [23]):

(8) a. Once there was **a wizard**, **he** lived in Africa, **he** went to China to get a lamp.

The choice of anaphoric means has been shown to adhere to discourse pragmatic principles, rooted in the concept of information flow, which captures how referents are introduced in discourse based on their information status [8, 9, 5, 23]. According to Chafe (1994), a referent's activation state can fall into three categories: new, accessible, or given. A new referent resides in the periphery of the hearer's consciousness; hence, it is less accessible, and must be introduced explicitly, e.g., with a full NP. An accessible referent is more active, possibly because present in the context or already mentioned, though not in the immediate discourse. In such cases, while a full noun phrase may still be required, an abbreviated version may suffice. In contrast, a given referent is already active, often due to recent mentions. Consequently, it is highly accessible, and can be encoded with a pronoun or a zero. Similarly, Givón argues that the choice among referential devices depends on the degree of referential or topic continuity, i.e., the possibility to identify the referent. The more continuous and predictable a discourse topic is, the more accessible it becomes, thus requiring less overt expressions.

(9) Referent-coding devices and referential continuity (Givón 1983)

> *lowest referential continuity*

a. indefinite NPs

b. definite NPs

c. stressed independent pronouns

d. unstressed anaphoric pronouns

e. zero anaphora

> *highest referential continuity*

The physical distance from the last mention of a referent contributes to determining the appropriate form. When the distance is short, a reduced form is more likely to be used, as in (8)

(a). Punctuation may also play a role. In (10), the unstressed pronoun in sentence (a) and the zero in sentence (b) both convey the idea of maximal referential continuity, thus suggesting the hearer to trace the referent back to the subject of the first clause, 'John'. However, sentence (b) represents an unsuitable continuation because, in English, zero anaphora can only be employed within chain-medial junctures, and not across chain boundaries [23].

- (10) John went to the mirror, \emptyset examined his hair, \emptyset sighed and \emptyset turned.
 a. Then he walked out.
 b. *Then \emptyset walked out.

Romance languages like Spanish and Italian were also observed to follow the hierarchy in (9); however, as morphologically rich, gender-system languages (cf. Section 7.2.1), they feature systematic pronoun subject dropping, obligatory pronominal agreement, and unstressed clitic object pronouns (11).

- (11) *Non* \emptyset_j *ho* *invit-ato* *Gianni_k* *perché* \emptyset_j *non* *lo_k* *cono-sco.*
 NEG have.PRS.1SG invite-PTCP because NEG 3SG.M.ACC know-PRS.1SG
 'I didn't invite Gianni because I don't know him.'

Inference-system, East-Asian languages like Chinese, on the other hand, make extensive use of zero anaphora [23, 45, 46]. What in English is done by pronouns, in Chinese is mostly done by zeroes. Using an overt pronoun instead of a zero may feel marked and instruct the reader to locate the antecedent in the distant text [42]. Zeroes may appear in different syntactic slots and cross both sentence and paragraph boundaries (12), also in cases of significant distance between the anaphor and its antecedent. Furthermore, in natural conversations, antecedents remain accessible even if separated by other intervening referents [45]:

- (12) (a) 片子取来了。(b) 孙主任看了, (c) 在场的医生都轮流看着。(Li 2005)
 (a) *piànzi_j* *qǔ-lái* *le.* (b) *Sūn* *zhǔrèn* *kàn* *le* \emptyset_j
 X-ray bring-over PFV/COS Sun dr. look PFV
 (c) *zàichǎng* *de* *yīsheng* *dōu* *lúnlú* *kàn* *zhe* \emptyset_j
 present DET doctor all turn look DUR
 'The X-ray was obtained. Dr. Sun examined (it). All the other doctors took turns to also examine (it).'

Both information flow and topic continuity theories offer a comprehensive framework for understanding reference management and have garnered support from cross-linguistic studies using natural discourse data. Psycholinguistic research has also related the use of referential devices to the cognitive aspects of memory and attention. Using a repeated name instead of a pronoun in consecutive sentences disrupts discourse coherence and prolongs processing time, a phenomenon known as the 'repeated name penalty' [25]. However, other research suggests that proximity and interference are not the sole factors influencing the choice of pronouns: in English discourse, too, a pronoun may be used even in the absence of an immediate antecedent [26]. In such cases, a shared common ground between the speaker and hearer can activate the reference, rendering the pronoun accessible.

7.2.3 Zero anaphora: challenges for parsers and NLP

Zero anaphora involves an empty grammatical slot representing a nominal referent in a sentence, devoid of lexical or grammatical markings. Despite its seemingly non-transparent nature, zero

anaphora is considered “one of the most natural, universal, ancient and functionally-coherent grammatical devices in the tool-kit of natural language” [24]. While zero anaphora is prevalent in all mature grammars, inference-system languages like Chinese, Korean, and Japanese exhibit even higher instances. Failing to detect and interpret zeroes in these languages can significantly impede comprehension.

(13) E. **He** dressed up, \emptyset_j went down the stairs_k, \emptyset_j walked out **the front door**_i and \emptyset_j closed **it**_i behind him.

I. \emptyset_j <i>Si</i>	<i>vest-ì,</i>	\emptyset_j	<i>scese</i>	<i>le scale</i> _k ,	\emptyset_j	<i>usc-ì</i>	<i>da-lla</i>
	REFL dress.up-PST.3SG		go.down-PST.3SG	the stairs		exit-PST.3SG	from.the
porta	d' ingresso _i	<i>e</i>	\emptyset_j	<i>la</i> _i	<i>chiu-se</i>	<i>a-ll'e sue</i>	<i>spalle.</i>
door	of front	and	3SG.F	close-PST.3SG	at-the	his	shoulder

C. 他穿好衣服，走下楼梯，走出前门后就顺手关上了。

$tā_j$	<i>chuān-hǎo</i>	<i>yīfú,</i>	\emptyset_j	<i>zǒu-xià</i>	<i>lóutī</i> _k ,			
3SG	dress-good	cloth		walk-down	stairs			
\emptyset_j	<i>zǒu-chū</i>	qián-mén _i	<i>hòu</i>	\emptyset_j	<i>jiù</i>	<i>shùnshǒu</i>	<i>guān-shàng-le</i>	\emptyset_i
	walk-exit	front-door	after	then	smoothly	close-up-PFV/COS		

The rules governing the licensing of zero anaphora show considerable variations among languages. In English, zeroes are permitted in subject positions when the subject corefers with that of the previous clauses, while objects must be overt (13) (E). In contrast, Italian requires the omission of all pronominal subjects without the need for prior mention; objects are obligatorily overt, either as stressed or as clitic pronouns, as *la* in (13) (I). In Chinese, any verbal argument can be represented as a zero as long as it encodes recoverable information (13) (C). Zero anaphors play a crucial role in maintaining discourse cohesion and contribute to the formation of topic chains [39].

As common as it may be, zero anaphora (ZA) remains a significant challenge for NLP. A first major issue is ZA detection. In the context of parsers, a probabilistic approach may be adopted, with many treebanks employing empty nodes in parse trees to represent ZAs [10, 4, 48]. Nevertheless, as this approach tends to limit the probabilistic engine from postulating the existence of zeroes everywhere, the representation of sentential meaning frequently becomes inconsistent [16, 17]. Specific challenges emerge with East-Asian languages, where ZA is prevalent. In the Chinese Treebank, zeroes appear once in every four sentences on average. This results in suboptimal degrees of accuracy levels in syntactic parsers, especially for subject/topic-drop languages like Italian and Chinese. More recent, deep learning based approaches accomplish ZA detection through two methods, i.e., syntactic analysis of sentences or binary classification of predicates, with varying degrees of accuracy [33, 29]. The second challenge is ZA Resolution (ZAR). While early studies used rule-based frameworks, with the increasing availability of annotated corpora, statistical and probabilistic models gained prominence. Later, various machine-learning models were developed, with recent trends leaning towards deep learning approaches that aim to automatically learn feature representations [37] Chen and Ng (2016) proposed a neural network-based approach incorporating both word embedding and hand-crafted features. Following studies, mainly focusing on East Asian languages, aimed to account for the long-term influence of antecedents [49, 30, 33, 44]. However, “the performances of best systems up to now are not high even after a long history of research” [33]. Despite being crucial for NLP applications, if anaphora resolution cannot be achieved with sufficient precision, it could pose more problems than it solves [37, 41] highlights outstanding issues regarding the performance limits of anaphora resolution algorithms and of knowledge-poor methods. Pre-processing accuracy remains suboptimal, affecting system performance. Most anaphora resolution systems require

manual pre-processing or pre-analysed corpora, and only a few operate fully automatically. Finally, Mitkov underscores the need for further research in multilingual contexts, which aligns with the objective of this paper (and this volume).

Partial coreference and bridging

A further challenge in machine representation involves partial coreference, also known as *bridging*. This phenomenon occurs when interpretation is based on a non-identical antecedent, hence relying on background assumptions or inference [13, 12, 28, 2] The bridging relationship may have different semantic facets, as illustrated below with examples in the cooking domain:

- (14) Lisa opened **the picnic basket**. **The sandwiches** were perfectly crafted. (Setting-element)
- (15) Cut **the lemon** and take **the seeds** out. (Whole-part/necessary parts)
- (16) Slice **the bread** into **pieces**. (Partitive)
- (17) Add all **dry ingredients**, except **yeast**. (Set-membership)
- (18) Add all **fruits**, except **strawberries**. (Hypernym-hyponym)
- (19) Bake for 15 minutes, then turn **the oven** off and let it cook. (Necessary role/tool)

Bridging cross-reference anaphora is most commonly encoded with definite NPs [28]. However, other instances are possible, such as with numerals, quantifiers, and measure words, and indefinite adjectives/pronouns, highlighted in bold in the following example from [20].

- (20) Remove sinew and membrane from the chicken breast, chop, or mince very finely to almost a paste. Mix with $\frac{1}{2}$ **a tablespoon of the corn flour**, milk [...] With one chopstick pick up **a little of the chicken mixture** and drop that in the stock, go on until **all the chicken** is used up and you have **lots of chicken bits/balls** pea size.

According to Clark (1977:413), bridging is an obligatory part of comprehension. The listener needs to identify reference for all referring expressions which, in most cases, requires "to bridge, to construct certain implicatures, [which are] a necessary part of comprehension". While this is easily accomplished by human brains, it poses significant challenges for computational technologies, which often struggle with handling partitivity and quantification expressions in anaphors [20, 19].

Evolving anaphora

Evolving anaphors are referential expressions referring to entities that, in the meantime, underwent some transformation, evolution, or gradual changes. These anaphors are frequently encountered in scientific, technical, and procedural texts, including recipes, but are also found in everyday language. Asher (2000) is credited as one of the first linguists to introduce the concept of evolving anaphora [1], even though its existence had previously been observed in recipe texts [7]. In (21), the anaphora *them* does not properly refer to *six cooking apples*, as they have undergone a change of state, i.e., they are now washed and cored.

- (21) Wash and core six cooking apples. Put **them** onto a fireproof dish. (Brown & Yule 1983: 201)

Sometimes, the antecedent is not overtly expressed, as in (22). Clearly, such anaphors are even more challenging for parsers to handle, especially when the anaphoric mean is a zero (23):

- (22) Add tomatoes and simmer for 30 minutes [...] Combine pasta and **sauce**.
 (23) Add tomatoes and simmer for 30 minutes [...] Combine **Ø** with pasta.

In 1983, Brown and Yule already emphasised the delicate nature of reference in this type of anaphora and proposed the need for a model of interpretation that allows textual referents to accumulate features or undergo changes as the discourse progresses:

The processor establishes a referent in his mental representation of the discourse and relates subsequent references to that referent back to his mental representation, rather than to the original verbal expression in the text [7].

Korzen devoted particular attention to evolving anaphora, which he defines as “the context in which a textual referent is referred to through an anaphoric expression after undergoing radical transformations in its ontological state” (2009:323, our translation):

- (24) *Lessate **le patate**, sbucciate-**le** e passate-**le** calde dallo straccio sopra a un velo di farina. Fate una buca sul monte **delle patate**, salate**le**, date **loro** l'odore della noce moscata [...].*
 ‘Boil the potatoes, peel them, and pass them hot from the tin over a layer of flour. Make a hole in the mound of potatoes, salt them, give them the smell of nutmeg [...].’

In (24), the potatoes designated by the antecedent *le patate*_j then become mashed. Despite undergoing a significant transformation, the referent is then recalled similarly with the PP *delle patate*_k ‘of the potatoes’. Korzen argues that this continuity arises from the reader recognizing a certain homogeneity [36]: although the entity experiences changes in its state, its main properties, such as colour, flavour, and smell, which directly influence the recipe, remain unchanged, as illustrated in Figure 1:

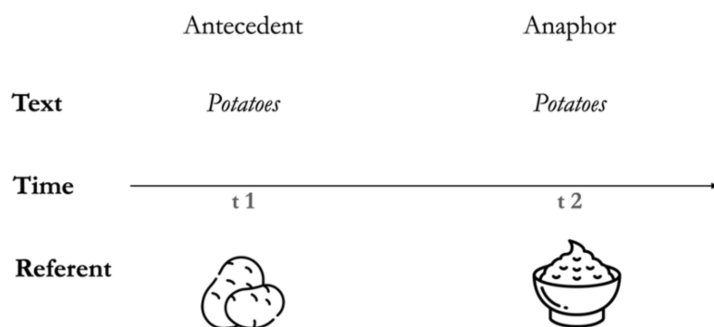


Figure 7.1: Figure 1 Evolving anaphora (elaborated from Korzen 2009)

Similarly, in (25), the referent of the noun *potatoes* undergoes significant changes in the various steps, although being referred to with the same noun throughout the text:

- (25) Bring a large pot of salted water to a boil. Add **potatoes** and garlic, lower heat to medium, and simmer until **potatoes** are tender, 15 to 20 minutes. When the **potatoes** are almost finished, heat milk and butter in a small saucepan over low heat until butter is melted. Drain **potatoes** and return to the pot. Slowly add warm milk mixture, blending it in with a potato masher or electric mixer until **potatoes** are smooth and creamy.

However, acceptability may vary based on the type of transformation and within certain limits, or *ontological thresholds* [36]. A critical threshold, for example, involves the change of state

from solid to liquid, which is so significant that a new form (the juice) is necessary (26) (Figure 2):

- (26) a. *Juice **the apples**, then put **them** onto a pan.
 b. Juice **the apples**, then put **the juice** onto a pan.

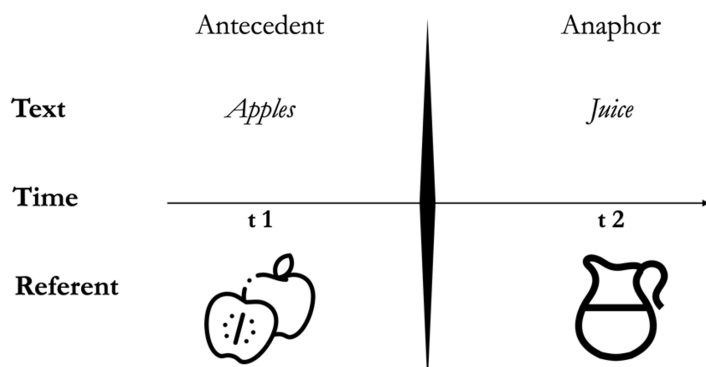


Figure 7.2: Figure 2 Evolving anaphora - threshold

Evolving anaphors constitute a significant challenge in NLP, especially when encoded through zero pronouns, as in (2) and (23). Challenges arise when multimodal anaphora resolution is involved, namely the process of understanding reference not only within the same modality (e.g., within a text) but also across different modalities (e.g., videos, or audios). This requires the model to comprehend and link information from diverse sources and identify *shared reference of entities*, namely when and how two modalities refer to the same description. In the context of cooking videos, the choice of a referring expression is influenced by both modalities. Oguz et al. (2022) analyse the discrepancies among in-text anaphoric means that are motivated by video materials: for example, the NP *the cubes* is used in a step to describe the bread pieces only because the instruction *chop the bread* shaped them into cubes in the video [43].

Recent progress has been made in addressing bridging, evolving anaphora and procedural text understanding. One prominent research direction involves representing procedures as workflow graphs structured around actions, with nodes representing named entities and edges representing relational structures, effectively capturing ingredient interactions. Yet, the approach falls short in addressing generic anaphora phenomena. Bosselut et al. (2018) predict action and entity attribute dynamics in recipes using a recurrent memory network, yet they do not directly tackle zero-anaphora [6]. The RISEC corpus [32] focuses on identifying zero-anaphora verbs in English recipes but relies on manual annotation and does not fully address general anaphoric phenomena. Some studies consider evolving anaphors as instances of bridging: Fang, Baldwin, and Verspoor (2022) propose an annotation framework for bridging anaphors applied to the RecipeRef corpus [19]. Still, as noted by Oguz et al. (2022), "temporally evolving entities present rich and, to date, understudied challenges for anaphora resolution". Research presented within this book also works towards this direction.

7.3 The study

Building upon the background laid out so far, this study endeavours to compare anaphoric phenomena in original recipes across English, Italian, and Chinese. After an analysis of the structure and flow of information in recipes, it seeks to identify the types and frequencies of

anaphoric devices employed in the texts, paying particular attention to the choice of different forms (e.g., full noun phrases over pronouns or zero forms). For the sake of the present work, a further distinction was made between endophora and exophora [27], although see Cornish (2010) for a more comprehensive taxonomy in terms of deixis and anaphora). Endophora refers to in-text reference relations, e.g. with elements present in the preceding text (like ingredients in the recipe) [15]. Exophora, on the contrary, points to entities existing in the context to which the utterance refers (e.g., utensils within the kitchen environment); it occurs primarily in texts where non-linguistic factors, such as the situational context, play a dominant role in semantic terms, with language serving as a secondary component, like recipes (1976, 24). Additionally, the study examines the nature of anaphoric relations, distinguishing between exact or partial coreference, evolving anaphors, and other relevant categories. The analysis is ultimately aimed at elucidating how the disambiguation process is allowed and assessing the extent to which inference is necessary.

As in Morbiato and Cani (this volume), the dataset comprises recipes describing the preparation of the same dish, pumpkin bread rolls, in three different languages: English, Italian, and Chinese. Authentic recipes were sourced directly from blogs (somewhatsimple.com, pandipane.blogspot.com, meishichina.com) while translations were avoided as they may reflect structures and word choice of the source language. Selected recipes were not edited by professionals and may present occasional grammatical imprecisions. Opting for user-generated recipes, as opposed to professionally crafted ones aims at filling the lack of research on contents freely produced by amateurs online, which display an interesting variety that professionally compiled recipes do not show (Morbiato and Cani, this volume). The choice of bread rolls recipes is deliberate: the complexity and diversity of steps (involving steaming, mashing, mixing, kneading, etc.) significantly enhanced the likelihood of identifying and comparing anaphoric means, in particular evolving anaphors. As typical blog entries, the texts are also correlated with pictures depicting each step, additional paragraphs that present the dish, a detailed list of ingredients and blog-like discussion about the procedure. Hence, multiple modalities are involved: the understanding of procedures is facilitated by both pictures and non-procedural textual parts.

The recipes were analysed by two independent raters who performed separate assessment and analysis of anaphoric forms in the data, in order to eliminate any potential bias. Referents and referential devices were highlighted in bold. Subsequently, interrater reliability was measured through a Cohen's Kappa test, which provided an assessment of raters' agreement of 0.995, indicating an almost perfect agreement between the two raters. The combination of independent assessment and the application of Cohen's Kappa test strengthened the reliability and validity of our research findings.

The current research was not specifically designed to achieve high generalisability, as the data corpus is relatively small. Instead, its objective was to gather initial qualitative data enabling cross-linguistic comparisons among typologically distant languages. Future research directions include a quantitative examination of larger corpora of recipes and procedural texts in general, to address the complex variations inherent in this textual genre.

Information flow and anaphoric choice in recipe texts

As observed, anaphoric coreference in recipes is not transparent, heavily relying on zero anaphors rather than explicit nouns. The agent, i.e., who performs all actions and processes, is typically left unspecified. Tracking coreference for patients, the entities undergoing these actions and processes, is also complex. In recipes, both bridging and evolving anaphors are common as ingredients (and specific parts or quantities of them) undergo swift and substantial changes. This leads to the creation of new entities, i.e., resultant objects, which, in turn, undergo further

movements and transformations (cf. Morbiato and Cani, this volume). Effectively addressing the intricacies of reference tracking in recipes requires breaking down complexity into more manageable components. To start, let us examine how information is organised within recipes.

The structure of recipes significantly influences the flow of information in the text, thereby impacting the choice of anaphoric means. A typical recipe consists of two essential components: a comprehensive list of the required ingredients and a detailed set of instructions outlining the steps for food preparation. The instructions primarily focus on the actions needed to transform the ingredients into a finished product and how to use the necessary kitchen tools for the task. In this component, ingredients are activated and highly accessible in the reader’s mind, whereas the new, most salient information revolves around the action and its result. Crucially, within the instructional compound, items may bear either exophoric reference, pointing to objects in the situational context—e.g., large bowl in (27), or endophoric reference, referring to above-listed ingredients (e.g., yeast, brown sugar etc.). At the beginning of the instructional compound, patients are generally chosen within the list of ingredients. As a consequence of the actions expressed by verbs, resultant objects (ROs) are created (Morbiato and Cani, this volume):

- (27) (E. 69) In a **large bowl** or **the bowl of a stand mixer** (EXOPH.), dissolve the **yeast in the warm water with a pinch of brown sugar**. (ENDOPH.: ingredients; semantic role: P)
 (E.70) Let \emptyset sit five minutes, until \emptyset foamy. (\emptyset : ENDOPH.: ingredients; semantic role: RO, then P)
 (E.71) Add the **remaining brown sugar, salt, pumpkin pie spice, egg, egg yolk, pumpkin puree, and 1/4 cup butter**, and mix \emptyset to combine. (ENDOPH.: ingredients; semantic role: P; \emptyset RO, then P)
 (E.72) Stir in **two cups of flour**, and mix \emptyset to combine. (ENDOPH.: ingredients; semantic role: P; \emptyset RO, then P)
 (E.73) Add **additional flour**, 1/4 cup at a time until the **dough** comes together and \emptyset is only slightly sticky to the touch. (ENDOPH.: ingredients; semantic role: P; **dough** RO)

As the recipe progresses, the resultant object of a specific step is the patient of the action in the following step—e.g., the frequent zeroes in (27) above. This is exemplified in Figure 3:

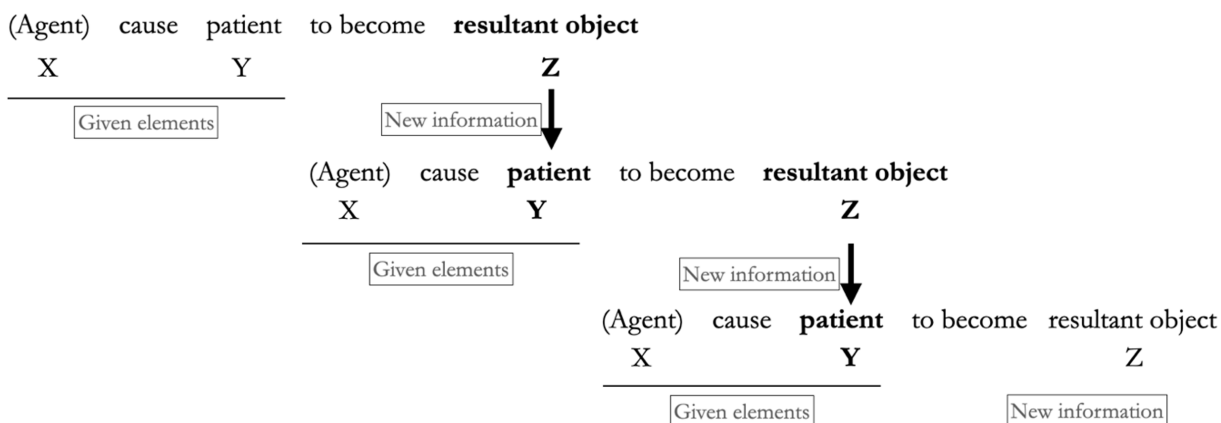


Figure 7.3: Figure 3: Argument structure and information flow in recipe texts

Crucially, this also captures how information flow progresses: the new, most salient part of each step resides in the action and its resultant object which, then, becomes given, highly active information in the next step. The pattern generally follows the sentence-initial given to sentence-final new transition. Processed entities (resultant objects) may also be recalled far later in the text, in a less continuous fashion. This happens often in modular recipes, as e.g.,

in recipes where a cake, cake filling, and cake icing are separately prepared, and only combined in a final step. As we have seen, theories of information flow and anaphoric choice predict that the activation state of referents influences the choice between more or less overt anaphoric means. In the following section, we will seek to determine whether this applies in our corpus as well.

7.3.1 Types of anaphoric devices

Table 1 shows the anaphoric devices employed in each language in the data. As all recipes involve the use of subjectless imperative sentences, the table does not include zeroes in the subject position. Full NP repetition is frequent in all three languages and so is zero anaphora, especially in Chinese, where pronoun use is absent. PPs are mainly used to introduce exophoric reference, specifically tools that a kitchen is supposed to be equipped with (e.g., *in a large bowl*, *into a blender*, *in forno* 'in the oven'). In Chinese, exophoric means may consist of nouns plus localizers, as *chúshījī-nèi* 'mixer-in' in (C.10), or also bare NPs, as *zhàzhījī* 'juicer' in (C9) (b) below. In any case, exophora is always expressed with overt means, be them NPs or PPs. Endophora, and specifically reference to ingredients and resultant objects, is encoded both through full repetition and weaker anaphors, although with significant differences among languages as illustrated in Table 1. Overall, Chinese makes the most extensive use of zero anaphora which, however, is very frequent in English, as well. Italian, on the other hand, maxes both use of zeroes and of clitics.

Referential devices	English	Italian	Chinese
Full NPs	33	25	18
PPs/nouns + localizers	14	20	7
Stressed Pronouns	2		
Unstressed/clitic pronouns		8	
Zero anaphors	16	11	28
Word count	329	321	411 (characters)

Figure 7.4: Table 1: Distribution of anaphoric devices

English employs full NPs, and to a lesser extent zero anaphora, both for patient arguments and for resultant objects—e.g., (E.74) and (E.76)—while pronouns are scarce. Notably, repetition of full NPs is used regardless of referents' activation stages, almost on par with zero anaphora. This is evident in (E.76), where the noun *dough* occurs twice despite it being highly activated and encoding exact coreference. Zero elements are typically recoverable from the text anaphorically, although some use of cataphora is also present. For instance, in (E.80), the zero (\emptyset) refers to 'rolls' mentioned later on in (E.82).

- (28) (E.74) Knead \emptyset 5-10 minutes, until \emptyset smooth.
 (E.75) Place \emptyset in a greased bowl and let \emptyset rise one, or until \emptyset doubled.
 (E.76) When the **dough** has risen, gently punch the **dough** down and divide \emptyset into two equal sized **balls**.
 (E.77) Roll each **ball** out into a **circle** that is approximately 18 inches in diameter and spread \emptyset with two tablespoons of the remaining butter.
 (E.78) Using a pizza cutter, cut the **dough** into twelve equal sized **wedges**.
 (E.79) Roll up \emptyset , starting from the wide edge, into a **crescent shape**.
 (E.80) Place \emptyset on a greased or parchment lined baking sheet, and repeat with remaining **dough**.
 (E.81) Roll the second ball of **dough** out, and repeat the process.
 (E.82) You should be able to fit all **24 rolls** on one large baking sheet.

In Italian, coreference is primarily encoded through full NPs, either repeated as, e.g., the word *impasto* ‘dough’ in (I.16), (I.17), and (I.19), or through synonyms/hypernyms: *polpa di zucca* ‘pumpkin pulp’ in (I.14) points to the same referent as *purea* ‘pumpkin puree’, which was used at the beginning of the procedure (*Prepar-ate la purea di zucca* ‘Prepare the pumpkin puree’). As expected, direct object deletion is present—e.g., (I.17), (I.18) and (I.19). Given referents in object position are also encoded through weak, clitic pronouns, e.g., *mettete-la* ‘put-it’ in (I.17) (b) and *sgonfiando-lo* ‘deflating-it’ in (I.19) (b). Resultant objects are expressed through NPs like *un composto cremoso*, lit. ‘a creamy mixture’ (I.14) (c) or *impasto* ‘dough’ (I.16.) (b), but also with zeroes.

- (29) (I.14)(a) *Mett-ete in una ciotola il quantitativo di polpa di zucca* (120gr),
 put-IMP.2PL in a bowl the quantity of pulp of pumpkin
 (b) *la Pasta Madre (a pezzetti), lo zucchero e parte d-ell' acqua* (es. 100 gr),
 the sourdough in pieces the sugar and part of-the water
 (c) *ora mescol-ate bene il tutto ottenendo un composto cremoso*.
 now mix-IMP.2PL well the everything obtaining a mixture creamy
 ‘Place 120 grams of pumpkin puree, the sourdough cut into pieces, sugar, and a portion of the water (e.g., 100 grams) in the bowl. Combine all ingredients thoroughly until you achieve a creamy texture.’

- (I.15)(a) *Aggiung-ete la farina e gradualmente inseri-te la restante acqua*
 add-IMP.2PL the flour and gradually insert-IMP.2PL the remaining water
 (b) *può darsi che non serva tutt-a*
 might be that NEG need all-F.SG
 ‘Add the flour and gradually incorporate the remaining water; you might not need all of it.’
 (c) ... (d) ...

- (I.16)(a) *Quando \emptyset sarà incordato aggiung-ete a filo l' olio di soia,*
 when \emptyset be blended add-IMP.2PL at thread the oil of soy
 (b) *impast-ate fino.a.che l' impasto non sarà omogeneo, liscio ed elastico.*
 knead-IMP.2PL until thedoughNEG be.FUT homogeneous smooth and elastic
 ‘When the dough is blended, slowly add some soy oil. Knead until the dough becomes homogeneous, smooth, and elastic.’

- (I.17) (a) *Pirl-ate un po' l' impasto a formare un-a bell-a palla lisci-a*
 spin-IMP.2PL a bit the dough to form a-F nice-F ball.F smooth-F
 (b) *ole-ate \emptyset leggermente e mett-ete-la in una ciotola ...*
 oil-IMP.2PL \emptyset slightly and put-IMP.2PL-F.3SG in a bowl ...
 ‘Give the dough a gentle spin to form a nice, smooth ball. Lightly coat with oil, place in a bowl and cover.’

- (I.19)(a) *Tolt-o* *da-l* *frigo* *fa-te* *tornare* \emptyset *a temp. ambiente per un' ora,*
 removed-M.SG from-the fridge make-IMP.2PL return \emptyset at temp. room for one hour
- (b) *stend-ete* *con il mattarello* *l' impasto* *sgonfiando-lo* *bene ma delicatamente,*
 spread-IMP.2PL with the rolling pin the dough deflating-3M.SG well but gently
- (c) *tirate* *l' impasto* *a-ll' altezza* *di circa 5 mm*
 stretch-IMP.2PL the dough at-the height of about 5 mm
- 'Take the dough out of the fridge and allow it to come to room temperature. Roll out the dough with a rolling pin, gently deflating it, and stretch it until it reaches around 5 mm in height.'

As expected in a gender-system language, anaphoric resolution is mainly achieved through clitic pronouns and gender/number agreement. For example, in sentence (I.19) (a), the reference of the masculine past participle verb *tolt-o* 'removed' (masculine) is *impasto* (masculine), which is mentioned cataphorically in the following clause, and not the closest NP *palla* (I.17), which is feminine and linked to the clitic *-la* in *mett-ete-la* 'put it' (I.17) (b). Yet, inference is needed, e.g., the NP *tutta* 'entire amount of' in sentence (I.15) refers to the word *acqua* 'water', and not to the word *farina* 'flour', both of which are plausible candidates (both are feminine and highly accessible). Correct disambiguation is achieved through common sense and world knowledge.

In Chinese, zero anaphors take precedence over pronouns and full NPs. Typically, resultant objects are subsequently referred to anaphorically through zeroes, with their coreferential scope often extending across sentence boundaries (topic chain). An antecedent might be introduced overtly and then recalled through zeroes, e.g., *kuài* 'pieces (of pumpkin)' in (C.9) (a), which later serves as the patient in the subsequent predicates *zhēng-shú* 'steam', *rù zhàzhījī* 'insert in juicer' and *zhà-chéng zhī* 'juice'. In other cases, the antecedent may be entirely absent, as seen with zero anaphors in sentences (C.12) to (C.15).

- (30) (C.9) (a) 南瓜泥，用南瓜去皮切块 \emptyset 蒸熟，(b) 然后 \emptyset 入榨汁机 \emptyset 榨成汁，

nánguā ní *yòng nánguā qù pí qiē kuài* \emptyset *zhēng-shú*
 pumpkin puree TOP use pumpkin remove skin cut pieces \emptyset steam-cook
ránhòu \emptyset *rù zhàzhījī* \emptyset *zhà-chéng* | *zhī*
 then \emptyset insert juicer \emptyset press-become juice

'For the pumpkin puree, peel the pumpkin, cut into pieces and steam, then put in the juicer and juice.'

...

- (C.12) 加入黄油，在继续启动搅拌程序， \emptyset 搅拌 5 分钟。

jiārù huángyóu *zài jìxù qǐdòng jiǎobàn chéngxù* \emptyset *jiǎobàn* *5 fēnzōng*
 add butter again continue activate mixing program \emptyset mix minute
 'Add the butter and continue mixing for 5 minutes.'

...

- (C.15) 然后 \emptyset 整理成面团，入不锈钢碗里。

ránhòu \emptyset *zhěnglǐ-chéng* *miàntuán* *rù* *bùxiù* *gāng wǎn* *lǐ*
 then \emptyset arrange-become dough insert stainless steel bowl in
 'Next, knead the mixture until it forms a dough and place it in a stainless-steel bowl.'

....

In Chinese, zero anaphora emerges as the predominant means to convey givenness and maintain topic continuity. Topic chains, such as that in (C.9), serve a functional role in recipes by presenting the sequential procedure more effectively and concisely. This allows the reader to concentrate on the patient and its transformations. The role played by inference in anaphora

resolution, as expected, is essential. Instances of ambiguous references requiring world knowledge and contextual tracking abound. For example, in (C.15) two full NPs are mentioned: *miàntuán* ‘dough’ and *bùxiù-gāng-wǎn* ‘stainless steel bowl’. Then, five clauses intervene, discussing other matters. In (C.19), a zero occurs as the argument of the verb *fājiào* ‘ferment’: it is world knowledge that tells us that the antecedent is the dough (which can ferment) and not the bowl.

7.3.2 Partial coreference and evolving anaphors

In all three recipes, various instances of partial coreference are present. In English, examples include NPs introduced by numbers, measure words, and adjectives indicating partitivity (e.g., *1/4 inch of water*, *the remaining brown sugar*, *two tablespoons of the remaining butter*) or else with nouns in a whole-part relationship, as e.g., *the insides of the pumpkin*. In Italian, in addition to the means mentioned above (e.g., *un dito d’acqua* ‘a finger of water’, *parte dell’acqua* ‘part of the water’), clitic partitive pronouns also occur, as e.g., ‘ne’, followed by the quantity–*usate-ne 120gr.* ‘use 120 g (of it)’ in (I.12). Hypernyms and collective nouns/expressions are used as well, such as *tutto* ‘everything in (I.14) (c) to denote all the aforementioned ingredients.

As expected, several instances of evolving anaphora are present in all three texts and are encoded by all sorts of forms (full NPs, pronouns, and zeroes). In the English recipe, the noun *pumpkins* in (E.22) indicates full uncooked pumpkins, while *pumpkin* in (E.29) indicates soft, cooked pumpkin. All mentions in between steps involve some change of state, such as cutting, meat scooping, and baking. Although ontologically different, these evolving entities are encoded by the same full NP, ‘pumpkins,’ or, alternatively, by pronouns (*them*, *they*) and zero anaphors. The author changes the anaphoric device to ‘puree’ only in (E.30) when the soft pulp is detached from the rind in (E.29). This can be illustrated through Figure 4.

- (31) (E.22) Cut the tops off your **pumpkins**
 (E.23) Cut the **pumpkins** in half, from top to bottom
 (E.24) Scoop out the insides of the **pumpkins**
 (E.25) Place the **pumpkins** face down on a baking sheet
 (E.26) Add about 1/4 inch of water around **them**
 (E.27) Bake \emptyset at 350 for 60-90 minutes, checking \emptyset periodically after 60 minutes
 (E.28) **They** are finished when a fork slides in easily
 (E.29) Let the **pumpkin** cool and then scoop the **meat** out into a blender.
 (E.30) Notes: If you want the **puree** to be more like what you get out of the can, drain the **puree** in a colander set over a pot. I found that paper towels worked great for lining the colander.

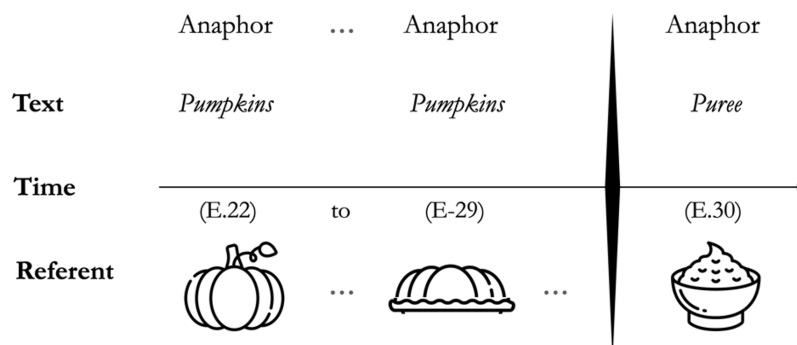


Figure 7.5: Figure 4: Evolving anaphora in the English recipe

Another example is provided in sentences (E.76) to (E.81) in (28), where *dough* refers to three distinct entities. In (E.76), it refers to the outcome of the amalgamation of the ingredients, the kneading, and the rising. In (E.78), it represents a flattened dough circle resulting from a ball that constitutes one half of the entire dough quantity. In (E.80) and (E.81), conversely, *dough* refers to the other half of the total amount of dough. Evolving anaphors are encoded through zero anaphors, as well.

Evolving anaphors are observable in the Italian and Chinese recipes, as well, with similar referential mechanisms. For Italian, one instance is (I.16)- (I.19) in (29), where the same referent undergoes several changes of states (in terms of temperature, thickness etc.), while being referred to with the same NP (*l'impasto*), clitic pronouns, or zero anaphors. In Chinese, evolving anaphors abound and are mainly encoded through zeroes; moreover, the antecedent may be missing altogether. For instance, zero anaphors in sentences (C.12) to (C.15) in (30) refer to the mixture obtained by combining all ingredients (\emptyset), which is never articulated as a full NP, but instead inferred contextually. Only in (C.15) does the author feel the need to specify that the mixture has now become a *miàntuán* ‘dough’. This can be exemplified with a figure similar to Figure 4.

7.4 Concluding remarks

This paper has highlighted the complexities of anaphoric phenomena in recipes. In terms of anaphoric means choice, both English and Italian employ a combination of overt means and zeroes which, crucially, also appear in object position, a phenomenon seldom encountered in other linguistic genres. Italian, however, also encodes highly activated objects with clitic pronouns displaying gender/number agreement. Conversely, Chinese exhibits greater reliance on zero anaphora wherein, in certain instances, overt antecedents are absent. Despite this, antecedent retrieval is allowed thanks to inferential processes/world knowledge. Importantly, inference proves indispensable not only Chinese but also in Italian and English, confirming its crucial role in anaphora resolution across languages. Any AI systems aiming to effectively address anaphora resolution and achieve complete textual comprehension must integrate inferential processes at a certain stage.

Crucially, the study has also uncovered instances where anaphoric means are selected irrespective of the activation status of antecedents, contrary to what predicted by theories outlined in Section 7.2.2. Repetition is observed also within the same sentence, i.e., with highly activated referents. Additionally, zero anaphora is identified in cases where the referent is relatively inactive. This latter phenomenon is not exclusive to Chinese, but also manifests in English, a language that is generally less tolerant in this respect. These findings challenge established expectations and highlight the nuanced and varied ways in which anaphoric reference operates in this genre. The use of both repetition and zero anaphora can be attributed to the dual nature of recipe texts. On one hand, they exhibit characteristics typical of procedural discourse, favouring precise reference encoding to avoid ambiguity. On the other, their colloquial register renders them conducive to the use of synonyms, pronouns, and a general avoidance of repetition. Recipes blend the need for clarity in conveying step-by-step instructions with the informality and flexibility inherent in everyday language use. Finally, instances of exophora, partial coreference (involving measure words, quantifiers, and whole-part relationships), and evolving anaphors abound in all three languages. Crucially, evolving anaphors are encoded both through overt means (e.g., full NPs and pronouns) but also with zeroes, which renders reference tracking more complex. Grasping the nuances of bridging and evolving anaphora requires a keen understanding of how references evolve over the course of instructions.

Chapter 7: References

- [CB1] Nichola Asher. Qualitative reasoning: Modeling and simulation with incomplete knowledge. In *Speaking of Events*. Oxford University Press, Oxford, UK, 2000.
- [CB2] NICHOLAS ASHER and ALEX LASCARIDES. Bridging. *Journal of Semantics*, 15(1):83–113, 02 1998.
- [CB3] Balthasar Bickel. *Grammatical Relations Typology*. Oxford University Press, November 2011.
- [CB4] Ann Bies, Mark Ferguson, Karen Katz, and Robert Mac-Intyre. *Bracketing Guidelines for Treebank II Style*. 1995.
- [CB5] John W Bois. Beyond definiteness: The trace of identity in discourse. ” the pear stories: Cognitive. *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*, 3:203–274, 1980.
- [CB6] Antoine Bosselut, Omer Levy, Ari Holtzman, Corin Ennis, Dieter Fox, and Yejin Choi. Simulating action dynamics with neural process networks. In *6th International Conference on Learning Representations (ICLR 2018)*, page 10, 2018.
- [CB7] Gillian Brown and George Yule. *Discourse Analysis*. Cambridge University Press, Cambridge, MA, USA, 1994.
- [CB8] Wallace L Chafe. Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In Charles Li, editor, *Subject and Topic*, pages 25–55. Academic Press, New York, 1976.
- [CB9] Wallace L Chafe. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. The University of Chicago Press, 1994.
- [CB10] Tagyoung Chung and Daniel Gildea. Effects of empty categories on machine translation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 636–645. 2010.
- [CB11] H H Clark and S E Haviland. Comprehension and the Given-New contract. social sciences working paper. *Social Sciences Working Paper. School of Social Sciences*, 1975.
- [CB12] H H Clark and S E Haviland. Comprehension and the given-new contract. Technical report, School of Social Sciences, University of California, Irvine., 1978.
- [CB13] H.H. Clark. Bridging. In *Thinking: Readings in Cognitive Science*. Cambridge University Press, Cambridge, 1977.
- [CB14] Bernard Comrie. *Some General Properties of Reference-Tracking Systems*. 1989.

- [CB15] Francis Cornish. Anaphora: Text-based or discourse-dependent?: Functionalist vs. formalist accounts. *Functions of Language*, 17:207–241, 12 2010.
- [CB16] Rodolfo Delmonte. *Computational Linguistic Text Processing - Lexicon, Grammar, Parsing and Anaphora Resolution*. Nova Science Publ, New York, 2009.
- [CB17] Rodolfo Delmonte. Predicate argument structures for information extraction from dependency representations: Null elements are missing. In *Distributed Systems and Applications of Information Filtering and Retrieval*, Studies in computational intelligence, pages 25–49. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.
- [CB18] Teun A Dijk and Walter Van. *Strategies of Discourse Comprehension*. Academic, New York, 1983.
- [CB19] Biaoyan Fang, Timothy Baldwin, and Karin Verspoor. What does it take to bake a cake? the RecipeRef corpus and anaphora resolution in procedural text. In Smaranda Muresan, Preslav Nakov, and Aline Villavicencio, editors, *Findings of the Association for Computational Linguistics: ACL 2022*, pages 3481–3495, Dublin, Ireland, May 2022. Association for Computational Linguistics.
- [CB20] Biaoyan Fang, Christian Druckenbrodt, Saber A. Akhondi, Camilo Thorne, Timothy Baldwin, and Karin Verspoor. Reciperef corpus for modeling anaphora resolution from the procedural text of recipes, url = <https://doi.org/10.17632/rcyskfvdv7>, howpublished = Mendeley Data.
- [CB21] William A Foley and van Valin. *Functional syntax and universal grammar*. Cambridge University Press, Cambridge, 1984.
- [CB22] S Garrod, D Freudenthal, and E Boyle. The role of different types of anaphor in the on-line resolution of sentences in a discourse. *J. Mem. Lang.*, 33(1):39–68, February 1994.
- [CB23] Talmy Givón. Topic continuity in discourse: A quantitative cross-language study. *Language (Baltim.)*, 63(1):160, March 1987.
- [CB24] Talmy Givón. *The story of zero*. John Benjamins, Amsterdam, 2017.
- [CB25] Peter C Gordon, Barbara J Grosz, and Laura A Gilliom. Pronouns, names, and the centering of attention in discourse. *Cogn. Sci.*, 17(3):311–347, July 1993.
- [CB26] S B Greene, R J Gerrig, G Mckoon, and R Ratcliff. Unheralded pronouns and management by common ground. *J. Mem. Lang.*, 33(4):511–526, August 1994.
- [CB27] M A K Halliday and Ruqaiya Hasan. *Cohesion in English*. Routledge, London, England, April 2015.
- [CB28] Yan Huang. *Anaphora: A Cross-Linguistic Study*. Oxford University Press, Oxford, England, 2000.
- [CB29] Ryu Iida and Massimo Poesio. A Cross-Lingual ILP solution to zero anaphora resolution. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 804–813. Association for Computational Linguistics, Portland, Oregon, USA, 2011.
- [CB30] Sei Iwata, Taro Watanabe, and Masaaki Nagata. Zero pronouns identification based on span prediction. In Jad Kabbara, Haitao Lin, Amandalynne Paullada, and Jannis Vamvas, editors, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language*

Processing: Student Research Workshop, pages 331–336, Online, August 2021. Association for Computational Linguistics.

- [CB31] Ray S. Jackendoff. *Semantic Structures*. MIT Press, Cambridge, MA, 1990.
- [CB32] Yiwei Jiang, Klim Zaporozhets, Johannes Deleu, Thomas Demeester, and Chris Davelder. Recipe instruction semantics corpus (RISeC): Resolving semantic structure and zero anaphora in recipes. In Kam-Fai Wong, Kevin Knight, and Hua Wu, editors, *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, pages 821–826. Association for Computational Linguistics, 2020.
- [CB33] Youngtae Kim, Dongyul Ra, and Soojong Lim. Zero-anaphora resolution in korean based on deep language representation model: BERT. *ETRI J.*, 43(2):299–312, April 2021.
- [CB34] Walter Kintsch and Teun A van Dijk. Toward a model of text comprehension and production. *Psychol. Rev.*, 85(5):363–394, September 1978.
- [CB35] Walter A Kintsch. The role of knowledge in discourse comprehension: A construction-integration model. In *Cognitive Modeling*, pages 5–48. The MIT Press, September 1988.
- [CB36] Iørn Korzen. Tipologia anaforica : Il caso della cosiddetta ‘anafora evolutiva. pages 323–357, 2009.
- [CB37] Kusum Lata, Pardeep Singh, and Kamlesh Dutta. A comprehensive review on feature set used for anaphora resolution. *Artif. Intell. Rev.*, 54(4):2917–3006, April 2021.
- [CB38] Stephen C. Levinson. *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge University Press, Cambridge, UK, 2000.
- [CB39] Wendan Li. *Topic Chains in Chinese: A Discourse Analysis and Application in Language Teaching*. Lincom Europa, München, 2005.
- [CB40] J M Mandler. *Stories, scripts, and scenes*. Distinguished Lecture Series. Psychology Press, London, England, January 2014.
- [CB41] Ruslan Mitkov. Anaphora resolution - the state of the art.” ———. 2022. “anaphora resolution.” in the oxford handbook of. In *The Oxford Handbook of Computational Linguistics*. Oxford University Press, Oxford and New York, 1999.
- [CB42] Anna Morbiato. Quello che i cinesi non dicono. In Magda Abbiati and Federico Greselin, editors, *Lingua Cinese: Variazioni Sul Tema (Sinica Venetiana 2)*, pages 79–102. 2015.
- [CB43] Cennet Oguz, Ivana Kruijff-Korbayova, Emmanuel Vincent, Pascal Denis, and Josef Van Genabith. Chop and change: Anaphora resolution in instructional cooking videos. In *2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing - Findings of the Association for Computational Linguistics: ACL-IJCNLP 2022*, pages 364–374. 2022.
- [CB44] Linfeng Song, Kun Xu, Yue Zhang, Jianshu Chen, and Dong Yu. ZPR2: Joint zero pronoun recovery and resolution using multi-task learning and BERT. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5429–5434, Online, July 2020. Association for Computational Linguistics.

- [CB45] Liang Tao. *Topic Discontinuity and Zero Anaphora in Chinese Discourse*. Amsterdam, Philadelphia: John Benjamins Publishing, 1996.
- [CB46] Liang Tao. Switch reference and zero anaphora: Emergent reference in discourse processing. In *Conceptual and Discourse Factors in Linguistic Structure*, pages 253—269. Stanford: CSLI Publications, 2001.
- [CB47] Wietske Vonk. The immediacy of inferences in the understanding of pronouns. In *Inferences in Text Processing*, page 29:205–18. 1985.
- [CB48] Nianwen Xue and Fei Xia. The bracketing guide- lines for the penn chinese treebank. Technical report, Technical Report IRCS-00-08., 2000. IRCS-00-08.
- [CB49] Qingyu Yin, Yu Zhang, Weinan Zhang, Ting Liu, and William Yang Wang. Deep reinforcement learning for chinese zero pronoun resolution. In Iryna Gurevych and Yusuke Miyao, editors, *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers*, pages 569–578. Association for Computational Linguistics, 2018.

