ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

Culture letterarie e filologiche

Ciclo 36

Settore Concorsuale: 10/G1 GLOTTOLOGIA E LINGUISTICA

Settore Scientifico Disciplinare: L-LIN/02 DIDATTICA DELLE LINGUE MODERNE

THE EXPRESSION OF DEFINITENESS IN ITALIAN NARRATIVES BY BILINGUAL CHILDREN: EXPERIMENTAL STUDY, AND TEACHING APPLICATIONS

Presentata da: Nicola Perugini

Coordinatore Dottorato
Marco Antonio Bazzocchi

Supervisore
Chiara Gianollo

Co-Supervisori
Sandro Caruana
Ana Teresa Pérez-Leroux

Esame finale anno 2024

A mio padre e mia madre

Riassunto

L’obiettivo di questo studio è analizzare l’espressione della definitezza in un compito narrativo in italiano somministrato a bambini bilingui e monolingui (di 4;0-8;3 anni). Sono stati esaminati tre gruppi di bambini bilingui, tutti con l'italiano come lingua maggioritaria e tre diverse lingue familiari: spagnolo, arabo e cinese. I sintagmi nominali sono stati estratti dalle produzioni narrative e annotati in base alla identificabilità del referente per l'interlocutore. È stata anche codificata la forma del determinante (definito, indefinito, dimostrativo o nome nudo). L’analisi mirava a verificare se la distribuzione delle forme in base all’identificabilità fosse la stessa in tutti i gruppi. I risultati hanno rivelato che tutti i bambini sono sensibili alla dimensione astratta dell’identificabilità del referente. I gruppi che parlano arabo e spagnolo come lingua familiare hanno mostrato una distribuzione dei determinanti simile a quella dei monolingui, mentre il gruppo che parla cinese è risultato differente, manifestando una preferenza per i nomi nudi o i dimostrativi quando il referente è identificabile. Questo è stato interpretato come un effetto della lingua familiare sull’italiano. Esaminando l’impatto di fattori interni (linguistici e cognitivi) ed esterni (quantità e qualità dell'input), è emersa un’interazione tra competenza morfosintattica e pragmatica nelle scelte referenziali. Fra i fattori esterni, le attività linguisticamente stimolanti e coinvolgenti svolte in italiano (es. giochi multimediali, giochi liberi coi compagni, lettura), sembrano avere un impatto positivo sulle abilità referenziali del bambino. Lo studio si è concluso con una proposta pedagogica che sperimenta gli effetti positivi della narrazione multilingue come strumento didattico in aule linguisticamente super-diverse.

Abstract

The aim of this study is to analyze the expression of definiteness in a narrative task in Italian administered to bilingual and monolingual children (aged 4;0-8;3 years). Three groups of bilingual children were examined, all with Italian as the societal language and three different family languages: Spanish, Arabic, and Chinese. Noun phrases (NPs) were extracted from narrative productions and coded based on whether the referent in the story was identifiable to the interlocutor. The form of the determiner (definite, indefinite, demonstrative, or bare noun) was also coded. The analysis investigated if the distribution of forms according to identifiability was the same across all groups. The results revealed that all children are sensitive to the abstract dimension of referent identifiability. The groups speaking Arabic and Spanish as their family language showed a distribution of determiners similar to monolinguals, while the Chinese-speaking group emerged as different, displaying a preference for bare nouns or demonstratives when the referent is identifiable. This was interpreted as an effect of the family language on Italian. Examining the impact of internal (linguistic and cognitive) and external (quantity and quality of input) factors, an interplay between morphosyntactic and pragmatic competence in referential choices was observed. Among external factors, linguistically challenging and engaging activities in Italian (e.g., multimedia games, free play with peers, reading) seem to have a positive impact on children’s referential skills. The study concluded with a pedagogical proposal experimenting the positive effects of multilingual narrative as a teaching tool in linguistically super-diverse classrooms.

CONTENTS

Introduction

The aim of this work is studying the expression of definiteness in Italian by three groups of bilingual children aged 4-8 years, all with Italian as the societal language and three different family languages: Arabic, Spanish and Chinese. Specifically, the expression of definiteness is investigated in the production of narratives, elicitated by means of a picture-based narrative task.

For a child, narrating a story is an engaging and highly significant activity due to the various benefits it provides for their linguistic development (Gagarina et al., 2016). However, becoming a competent storyteller can be a challenging task, as it requires the integration of diverse types of information—perceptual and linguistic—and the activation of various cognitive skills.

To become skilled storytellers, children must acquire the ability to effectively communicate sequential and causally connected story events, ensuring clarity for the listener regarding the who, what, and where of the narrative. This task entails the child introducing characters and entities in a manner that establishes enough common ground between the storyteller and the listener. Furthermore, during subsequent references to these characters within the narrative (maintenance and reintroduction), the storyteller needs to assess whether the referent is currently accessible to the listener and select a suitable linguistic form accordingly (Ariel, 2001; Gagarina & Bohnacker, 2022; Gagarina & Musan, 2020).

In this process, the speakers require language elements that designate the entities they intend to refer to. Linguistic expressions fulfilling this role are termed referential or referring expressions, encompassing proper names, natural kind terms (*grass*, *fire*, *lion*), indexicals (*me*, *you*), and definite descriptions (*the linguist*, *the sun*). In the following pages, particular attention will be devoted to definite descriptions and the determiners used to construct them. They can be studied from the perspective of grammaticality, i.e., the presence/absence of the determiner in obligatory contexts (resulting from language-specific syntactic-semantic restrictions), and its morphosyntactic correctness (e.g., agreement with the noun in gender and number). Additionally, they can be examined

from the standpoint of their discourse-pragmatic appropriateness, i.e., whether they reflect the information status of the referent and its accessibility to the listener. The latter will be the privileged perspective in this work.

The use of definite or indefinite expressions is said to be based on the distinction between new information and given information, as well as on the assessment of shared knowledge between the listener and the speaker. The information already given or shared should be marked as definite, while the new and unshared information should be marked as indefinite. For a pragmatically adequate use of definiteness markers, children must first develop sensitivity to such a distinction and be able to assess mutual knowledge (Clark & Marshall, 1981) and, consequently, adopt the perspective of their listener.

Within the array of theoretical frameworks addressing definiteness, this study adheres to the theoretical analysis of definiteness as identifiability as posited by Lyons (1999), among others. This account situates the phenomenon within the speaker-listener relationship since definiteness emerges as a consequence of the speaker's assessments regarding the interlocutor's knowledge state. It also provides an operationally convenient basis for implementing an annotation system of definiteness in narratives in which children's pragmatic competence can be assessed.

In order to adopt the perspective of the listener and assess mutual knowledge, acquiring sufficient Theory of Mind (ToM) skills is often said to be essential (Astington & Pelletier, 2005; Tomasello, 2003), but in addition, the child must also possess the necessary linguistic (lexical, syntactic, and morphological) abilities to use forms appropriately. Perspective-taking involves social and language skills, with evidence that it may also require additional executive function skills (De Cat, 2015). Moreover, the cognitive load varies based on the interaction or context, influencing referential expression production. Experimental tasks, like the picture-based narrative task in this study, require linguistic proficiency, attentional resources, and a well-developed working memory (Whitely & Colozzo, 2013). In sum, adequate reference in narrative discourse demands high processing capacities. The task may be especially challenging for bilinguals, who need to develop the

ability to produce adequate referents in both their languages. Moreover, bilingual children must choose the form according to the interlocutor's knowledge state, produce that form in a language-specific correct manner, and inhibit the options in the non-target language (Lindgren et al., 2022). Narrative has thus emerged as an optimal form of production for studying referential skills. Stories also serve as an effective educational tool for fostering the development of these skills in children. Becoming a competent narrator requires exposure to numerous stories and direct involvement in the production of narratives (Kerry-Moran & Aerila, 2019). This way, the child learns to create coherent and understandable narrative texts, where the perspective of the interlocutor is consistently taken into account, and the elements of the story are presented accordingly. More broadly, narrative tasks contribute to the overall linguistic development of the child, providing support not only in linguistic aspects but also in emotional and social development. These are the reasons why narration has been chosen as an educational proposal to present to teachers in the schools that have welcomed the data collection for the present research project.

Sustaining children's linguistic development can be particularly challenging for teachers working in schools with high levels of linguistic diversity, such as those found in contemporary Italy. Bilingual students with a migration background, are often exposed to less input in Italian and undergo a unique language learning trajectory. Narratives have also proven to be a valuable means to implement multilingual pedagogical practices aimed at enhancing the rich linguistic repertoire of highly diverse classes.

In sum, the study employs narratives both as a tool to investigate referential skills, specifically the expression of definiteness, and as an educational instrument to enhance these skills and, more broadly, the linguistic development of children.

Three research questions led the investigation:

1. Do bilinguals and monolinguals distribute articles differently according to identifiability?
2. Does family language influence the societal language in the acquisition of definiteness?
3. Is there an effect of child-external and internal factors on the acquisition of definiteness?

In order to answer the questions narrative data from bilingual children (10 for each linguistic group) and their monolingual peers who constituted the control group were collected. Having three different bilingual groups allowed us to distinguish the influence of the family language from more general effects of bilingualism on children's productions. To assess participants' linguistic dominance data were collected in both the family language and Italian. Narrative tasks from the Multilingual Assessment Instrument for Narratives (MAIN) (Gagarina et al., 2012) was utilized. The Italian data were then analyzed using a coding system based on definiteness-related features and quantitative analyses were conducted on these data. A questionnaire to gather information on children's linguistic experience was administered to both families and teachers.

Additionally, as mentioned, the project included an educational intervention component based on multilingual narratives as a tool for enhancing multilingualism and supporting the overall linguistic development of bilingual children to ensure harmonious bilingualism.

The work's presentation is organized according to this structure: in chapter 1 the account of definiteness and identifiability chosen for this study is presented, providing the theoretical framework for the subsequent research work. Chapter 2 describes the expression of definiteness in the languages involved in the study: Italian, Spanish, Arabic, and Chinese. In chapter 3 the acquisition of definiteness in monolingual and bilingual children is outlined, presenting the models used to interpret the data of this study, and to analyze the key factors influencing bilingual linguistic development. Chapter 4 introduces the participants and their recruitment, data collection methods and tools, and the coding system. In Chapter 5 the results of the work and their discussion are provided. Chapter 6 describes the multilingual narrative activities proposed to the classes from which the children were recruited. The work closes with conclusions.

1. Definiteness and identifiability

In this first chapter, the theoretical framework underpinning the research is presented to provide the foundations for whole work. The chosen theoretical framework for this work posits definiteness as the grammaticalization of identifiability and is one among various theoretical proposals developed to account for definiteness across languages. Definiteness has been at the center of linguistic debate since the late 19th century with the foundational works of Frege (1892) and Russell (1905). Since then, theoretical linguistics, philosophy of language, and, more recently, cognitive psychology have continuously generated proposals to capture salient aspects of this intricate phenomenon crucial to human communication.

The chapter is structured into four sections: section 1 describes the relationship between definiteness and identifiability, introducing some related theoretical concepts. Section 2 presents identifiability as a cognitive category and delineates its functions within the discourse universe. Section 3 introduces the continuum of identifiability and its linguistic manifestations. Section 4 provides an overview of the main formal means employed by languages to encode the category of identifiability.

1. Definiteness as the grammaticalization of identifiability

To achieve success in communication, cooperation between speaker and interlocutor is essential; the former guides the latter in understanding, by providing all the necessary elements for that purpose, the latter cooperates and accommodates. Among those elements guiding the hearer in navigating the discourse and linking text and extralinguistic elements, referential expressions play a key role. They serve to identify, in the discourse or in the external reality, the entities being spoken about. This creates a three-place relationship, in which speaker *x* uses referential expression *y* to refer to entity *z*

(Abbott, 2010). That is, referential expressions function as footholds in that they serve to ensure that the addressee has a cognitive fix on the same object that the speaker does (Bezuidenhout, 2019).

Critical in this respect is the distinction between referents that the hearer already knows and those that the hearer does not yet know. Evans and Levinson (2009:437) argue that there are "functional features that all languages need in order to be adequately expressive instruments" and among them they mention the ability to distinguish new from old information. The sharing of information between speaker and listener can arise from what has already been mentioned in previous discourse, but it can indeed also stem from what Turner (2021) defines as "shared cognitive perceptions of the world, entities within it, and other humans' knowledge of them" (Turner, 2021:2). Languages such as Italian or English have dedicated elements to encode the distinction between what the hearer already knows and what the hearer does not yet know, traditionally known as definite articles, if they are used to mark information as given or shared, or indefinite articles, if they are used to mark information as discourse-new or unshared. Other languages, on the other hand, such as Russian or Japanese, do not possess any kind of article. However, since the distinction between new and old, shared, and unshared information is fundamental to communicative success, we must assume that all languages have some means of conveying it, regardless of the presence of overt marking of definiteness.

A difference needs to be drawn between definiteness as a grammatical category and definiteness as a category of meaning. Lyons (1999) suggested that definiteness *stricto sensu* is a grammatical category like tense, mood, number, and gender. It is the representation in grammar of some category of meaning, just like those mentioned above. There is never a 1:1 relationship between a grammatical category and the category of meaning it is based on. Provided this, we can expect that grammatical definiteness is not found in all languages, like other grammatical categories. Defined in such a way, definiteness is present only in languages that exhibit an overt marking, like articles.

According to Lyons, grammatical definiteness is the grammaticalization[1] of "semantic/pragmatic definiteness".

Taking Löbner's (1985) differentiation between semantic definiteness and pragmatic definiteness as a foundation, we consider semantic definiteness reliant on the inherent semantic content of nouns. Examples of 'semantic definites' include proper nouns and nouns denoting unique entities such as the sun, the moon, the weather, and the air. These entities are considered definite by virtue of their inherent characteristics. In contrast, pragmatic definiteness is not tied to the type of nouns and is instead associated with the context (discourse universe, the perceptual context etc.) (see also Napoli, 2009).

Therefore, total equivalence between definite and old information, and indefinite and new information, cannot be asserted, as noted by Chafe (1976). The notions of definite/indefinite and given/new are not exactly overlapping and should be kept separate. Givenness is related to pragmatic definiteness, but not to semantic definiteness. In fact, a referent introduced for the first time in discourse may be easily identifiable by the hearer, albeit at first mention (i.e., discourse-new), because it is already present in his or her knowledge. See the following example from Chafe (1976:43):

(1) *I talked with* ***the carpenter*** *yesterday.* (Uttered at the beginning of a conversation)

The NP in bold is both definite and new, and in a case like this, definiteness is established on another basis than the immediate previous mention, which would have created givenness. The notion of given/new is therefore not sufficient to define the concept of semantic/pragmatic definiteness; other grounds for it must be sought.

[1] Lyons uses the term grammaticalization "essentially in a synchronic sense, to denote the representation by a grammatical form or forms (and thus with the status of a grammatical category) of some concept of meaning" (Lyons 1999: 276).

The debate on definiteness has been going on for more than a century and definitions are numerous; the one used in this chapter goes back to Chafe (1976) and Lambrecht (1994) and refers to the notion of identifiability.

Therefore, quoting Becker (2018), referents are definite if "they are identifiable by both the speaker and the hearer, which I refer to as mutual identifiability. To be more precise, the referents need to be unambiguously identifiable in order to be definite, since their unambiguous identifiability guarantees that they are the only salient referent of their kind" (Becker 2018:76).

1.1 Related notions: referentiality and specificity

Once the difference between grammatical definiteness (formal encoding) and semantic/pragmatic definiteness (category of meaning) has been defined, it is useful to clarify two notions closely related to definiteness: referentiality and specificity. Despite the lack of agreement among linguists on these notions and the heterogeneity of related terminologies, this section will attempt to briefly summarize only the fundamental issues that are pertinent to the goal of this work.

Linguistic referentiality is both a semantic concept and more substantially, a pragmatic concept. Linguistic expressions in certain formal categories, as noted in Lyons (2002:297), ''may be factorized, semantically if not syntactically and lexically, into two components. One of these is descriptive (e.g., the word '*man'* in '*the man'*); the other is purely referential (e.g., the definite article '*the'* in English)''. Semantic referentiality, which essentially serves the function of pointing to some existent entity in a model of discourse, defines the referential part of the expression. Semantic referentiality is a property that is encoded in the semantics of linguistic expressions. Some linguistic expressions are inherently referential (demonstratives, pronouns, and proper names), others inherently nonreferential, such as negative quantifiers and partitive quantifiers (such as 'each', 'every', 'all', 'both' and 'most'), and others are ambiguous in referentiality, which is the case with definites, indefinites, and bare nominals (Chen, 2009; Lyons, 1977; Partee, 1970).

If semantic referentiality is something that words do, pragmatic referentiality is something that speakers do, and it is context-dependent (Abbott & Gundel, 2019). Referring is something people use expressions to do; it sets up a three-place relation between speakers, linguistic expressions, and the world. It is concerned with how speakers, by using a particular expression (e.g., *she, that woman*), succeed in picking out their intended referent for the addressee, and how the use of different types of expressions in different contexts affects addressee interpretations (Abbott & Gundel, 2019).

This is the type of referentiality we will mainly refer to in this chapter. Pragmatic referentiality, being context-dependent, is a strongly discourse-dependent notion. In this work, we will follow Du Bois (1980) who provides the following definition of referential: "a noun phrase is referential when it is used to speak about an object as an object, with continuous identity over time" (1980:208). He provides the following examples:

(2) *And* ***a boy*** *comes by riding* ***a bicycle*** *… And he sees* ***the pears*** *… and he stops.*
(3) *he looks like* ***a uh .. Chicano American.***

In the first example, all the NPs are referential: '*a boy*' and '*a bicycle*' introduce a non-identifiable referent that is new and can be referred to later in the discourse; while '*the pears'* refers back to something that is already present in the previous discourse and in that is identifiable. The referential concept is bounded and may serve as a focus for future reference. This is what is meant by continuity of identity. Each NP can be followed by other referential phrases, that refer to the same referent.

In the second example, '*a Chicano American*' is a non-referential NP; in fact, the speaker is not talking about a Chicano American, but rather is attributing to a subject (*he*) the characteristics of a Chicano American. Predication (*he looks like*) serves to abstract these features from any concrete referent and assign only the descriptive content of the expression to the subject. In the following discourse '*Chicano American*' cannot be taken up anaphorically and encoded with a definite article (Du Bois, 1980). Trying to classify in a few groups all types of non-referential NPs reported in the

literature, Chen (2009) identifies five different non-referential usage patterns, encompassing all nominals that are used primarily for their descriptive content and not to refer to individuals:

(4) A. Generic

The tiger *is a large, fierce animal*. (Chen 2009: 1161)

B. Qualitative

John is ***a math teacher****.* (Chen 2009: 1161)

C. Idiomatic

She is on ***the phone*** *right now.* (Chen 2009: 1161)

D. Non-ostensive:

The murder ***of Smith*** *is insane!* (Donnellan, 1966:297)

E. Nonspecific

John intends to marry ***a Norwegian girl****. She must be a linguist.* (Heim 1988: 249)

These are the uses that, for the purposes of this work, are treated as non-referential and cover both formally definite and indefinite expressions. Expressions in the first three groups (4A, B, C) are nonreferential in that they do not presuppose the existence of an entity as described by the linguistic encoding of the expression. Among these, the first type of reference (4A) in Chen's classification is generic reference: instead of referring to a specific individual, it describes a kind or genus (see Carlson, 2011; Krifka et al. 1995, for a detailed discussion). The non-ostensive reference in (4D) is from Donnellan (1966) who characterizes the second member of the expression (*murder of Smith*) as attributive (label also used by Partee, 1970), meaning that by using it the speaker does not have the intention to draw the attention of the addressee to a particular referent. The example in (4E) is particularly interesting in our discussion since it introduces the concept of specificity, to which we turn.

Following Chen (2004), an expression is specific only if the speaker uses it to refer to a particular entity in the world of discourse, independently of its being identifiable or not by the addressee. While the concept of identifiability is hearer-oriented, in the sense that the speaker must take into account the knowledge of the addressee, the notion of specificity is speaker-centered (Chen, 2004), in the sense that "the speaker has a particular individual in mind" (Karttunen, 1968: 20)[2].

In languages with explicit identifiability marking, a non-identifiable referent will typically be encoded as non-definite, even when it is specific (that is, well-known to the speaker).

The specific/non-specific distinction is thus mainly applied to indefinite NPs, which are specific if the referent is clear in the speaker's mind, although the speaker knows that it is not identifiable by the hearer. A non-specific indefinite, on the other hand, is non-referential, so in that case the function of the article is no longer to mark an opposition of identifiability. Considering the following very famous two examples:

(5) a. *John intends to marry **a Norwegian girl**. She is a linguist.* (specific)

b. *John intends to marry **a Norwegian girl**. She must be a linguist.* (nonspecific)

In both cases, the hearer has no knowledge about the referent and is unable to identify it, but in example (5a) the speaker clearly refers to a specific person she/he knows and has in mind, while in (5b) she/he does not. The specific or non-specific reading does not emerge directly from the NP but is the result of interaction with other elements of the utterance, such as verbal aspect (see Givón, 1990 for a detailed analysis). As it can be seen from sentence (5b), the non-specific NP, although non-referential, admits of anaphoric recovery in subsequent speech (Heim, 1988). Having clarified the

[2] The notion of specificity has been extensively investigated in linguistic studies. It must be acknowledged that the view adopted here does not account for many other proposals in the literature (see Kartunnen, 1968; 1969/1976; Partee, 1970; Kripke, 1977; Fodor & Sag, 1982; Enç, 1991 among others). For a comprehensive overview of the main issues and theoretical contributions related to specificity, see von Heusinger (2019).

two notions of referentiality and specificity, let us return to definiteness and look at it in its discourse-pragmatic dimension.

2. Definiteness in discourse and mind

2.1 Identifiability as a cognitive category

In continuity with the work of Lambrecht (1994), we argue that when a speaker wants to make an assertion, involving some entity that he/she assumes is not yet represented in the mind of the hearer and cannot be referred to deictically, it is necessary for him/her to create a representation of that entity through a linguistic description, that can be anaphorically referred to in the subsequent discourse. The creation of this new mental representation for the addressee can be compared to the establishment of a new referential file in the discourse register,[3] to which further elements of information can be added and which can be reopened in the future[4]. The difference between entities for which the speaker can assume there is already an open file in the mind of the hearer, and entities for which a new file has to be opened, can be explained precisely by the category of identifiability. A referent is identifiable if a shared representation in the mind of the speaker and listener already exists at the time of utterance, while a nonidentifiable referent is one for which there is not yet a representation in the mind of the listener, but only in the mind of the speaker. Identifiability is thus configured, not only as a discourse-pragmatic category, but also as a cognitively based category. And it is precisely because of this nature that we can assume that the mental ability to identify referents is the same in speakers of all languages.

Whenever a nonidentifiable referent is introduced into the discourse for the first time, the speaker must signal the listener to open a new mental file. This referent will receive an indefinite formal encoding, which will be the signal for the establishment of the new file. Opening a new file tends to raise the expectation that the referent will continue to be used as more information is added

[3] Lambrecht defines the notion of discourse register as "the set of representations which a speaker and a hearer may be assumed to share in a given discourse." (Lambrecht 1994:74).
[4] On the metaphor of 'file' see also Heim (1982).

to it, but it is also possible that this expectation will not be met. If that file is to be reopened later, at the reopening, its identifiability will be signaled, and the formal encoding will be definite. In languages that have grammaticalized identifiability, it is up to the articles (or language specific dedicated elements) to signal to the listener whether a new file needs to be opened (Lambrecht, 1994). In other words, the article performs the function of indicating to the interlocutor if he/she is in the condition to identify the referent (Lyons, 1999).

From this it follows that (in)definite descriptions, like all referential expressions, activate specific mental operations in the interlocutor, which pertain to two cognitive domains:

- attentional activation
- memory: searching an existing file, retrieving from an existing file, adding input into an existing file, opening a new file.

In many discourse contexts, grammar activates these cognitive operations in both domains. Grammatical signals, such as articles, can thus be seen as mental processing instructions. Studies of grammar in text suggest that referential coherence grammar serves primarily to identify and activate mental files. Nominal referents come to be file labels, useful for identifying, accessing, and activating files in memory where the coming information is stored (Givón, 2017). Activating a file means that the referent enters the hearer's consciousness, i.e., the listener creates a mental representation of it, which sits in memory and can be retrieved (Chafe, 1976; Lambrecht, 1994). It is precisely the operations of activation, retrieval, and deactivation that create the continuity of identity and make it possible for the interlocutor to trace identity.

Let us now look at what happens specifically for definite and indefinite expressions, following the mental processing paths suggested by Givón (2017). The sequences (6) and (7) are reported and named as in Givón (2017:27, 29).

Givón (2017) takes the listener's perspective and in (6) and (7) proposes two flowcharts representing the cognitive processes undergone by an addressee following the mental instructions provided by the articles while navigating the text. The definiteness status of a referent determines

whether the file to be activated in hearer's mind is a new one (indefinite) or one that already exists (definite). It should be specified that in the model proposed by Givón, the activation of a file implies the deactivation of the currently active referent, but in the case of (in)definite descriptions (as well as other grammatical devices for encoding larger-size referents: stressed pronouns or proper names; see Section 1.2.2 below), this deactivation is not automatic. In fact, the first mental operation that a listener must perform is not related to the definiteness status of the referent (i.e., deciding whether to open a different file from the current one) but to decide whether the referent in question is important or not. By importance, Givón refers to topical referents (he also uses the term "thematic importance"; Givón, 2017:23), meaning those that have a continuity of mentions in the subsequent discourse[5]. Only important (topical) referents are activated and then serve as file labels for the incoming information. Referents that the grammar has designated as important continue to be discussed and remain cognitively active. In other words, the thematic importance of a new referent determines whether the currently active file should continue as an address for the incoming information or not. Once this decision is made, the hearer can proceed with operations related to definiteness and activation.

(6) Indefinite grammatical marker as mental processing instructions:

a. FULL NP ===> Defer activation decision; then
b. if INDEF ===> Do not search for an existing file; then
c. if UNIMPORTANT, then
(i) do not open a file
(ii) do not activate
(iii) file as new information in the current active file.
d. if IMPORTANT, then
(i) Open a new file in episodic memory;
(ii) activate;
(iii) start filing incoming information in new file.

(7) The grammar of definite reference as mental processing instructions:

a. If DEF NP ===> Defer activation decision; then
b. if UNIMPORTANT, then
(i) do not open a file
(ii) do not activate
(iii) file as a chunk of new information in the currently active file.
c. if IMPORTANT, then
d. Determine the source of definiteness among the disjunctive options
(i) situational ('deictic') working memory or attention
(ii) generic ('cultural') semantic memory
(iii) textual ('discourse') episodic memory; then
e. Search for antecedent co-referent in the appropriate mental file; if found, then
(i) retrieve;
(ii) re-activate;
(iii) Start filing incoming information in the re-activated file.

[5] For a discussion on thematic importance and how to measure it, see Givón (1983; 2017:13) and Wright & Givón (1987).

As it is evident from the processing sequences (6) and (7), any choice about the activation of both definite and indefinite NPs must be postponed until the referent's thematic importance/topicality has been established (6c or d; 7c or d).

The grammatical marker "indefinite" indicates that there is no need to search in the memory (6b). If there is a grammatical signal qualifying the NP as an "important indefinite" (6d), the hearer is told to open and activate a new file for the referent (examples of important indefinites include referents introduced for the first time through NPs such as 'this + noun' in spoken American English[6]).

Important definite NPs (7c), in addition, perform two additional crucial cognitive operations: finding the source of identifiability (7d) and searching the storage space for an existing file (7e). According to Givón (2017)'s proposal, this would make definites more complex than indefinites.

In the sequence (7), there would be three main types of sources from which the identifiability of the referent stems: speech situation, shared general knowledge and mentally represented text, which are related to three different cognitive domains (Atkinson & Shiffrin 1968): working memory/attention, lexical semantic memory, episodic memory. Accessing the first type of referents is a quick and automatic operation, as the concept in question is under attentional focus. Semantic and cultural memory is triggered directly by vocabulary; referents in episodic memory, on the other hand, are triggered by highly specific grammatical cues (Givón, 2017). Section 1.3.1 will be devoted to a more detailed examination of the various types of sources of identifiability that license the use of definite expressions.

2.2. (In) definite referents as discourse signals: accessibility and cognitive status

As just seen, during the unfolding discourse, for the hearer it is necessary to access the referent and its activated file. However, not all referents are equally accessible. Mira Ariel (1988, 1990, 1991,

[6] For a discussion on indefinite demonstratives, refer to Prince (1981), Wright & Givón (1987) and von Heusinger (2011) among others.

2001) proposes the Accessibility theory, according to which referential forms are determined by how accessible the referent is.

"Under the assumption that mental representations (specifically those of NPs) are accessible to addressees in varying degrees, the claim is that speakers choose between referring expressions so as to mark such accessibility differences for the addressee's convenience." (Ariel, 1991:443).

Ariel identifies a hierarchy of accessibility, ranging from low accessibility forms, through medium accessibility forms, to highly accessible forms. Referents will be encoded in forms that are the lighter the more accessible they are. More informative, less ambiguous, more marked, and longer forms retrieve less accessible referents. For example, full NPs + modifiers are signals of minimally accessible referents, while zero anaphora is a signal of maximal accessibility. The hierarchy from Ariel (2001: 31) is provided below:

(8) Full name + modifier > full name > **long definite description** > **short definite description** > last name > first name > distal demonstrative + modifier > proximate demonstrative + modifier > distal demonstrative + NP > proximate demonstrative + NP > distal demonstrative (-NP) > proximate demonstrative (-NP) > stressed pronoun + gesture > stressed pronoun > unstressed pronoun > cliticized pronoun > verbal person inflection > zero.

We will not describe here Ariel' entire theoretical proposal, which takes into consideration multiple factors that determine the accessibility of referents; what is of our interest is the place Ariel assigns to definite descriptions in this hierarchy. For the author, they would be markers of relatively low accessibility. It should be mentioned, however, that Ariel distinguishes between long (more than two content words) and short (one or two content words) definite descriptions, where the former signal lower accessibility than the latter (Ariel, 1991).

For “old” NPs, which evoke a representation already introduced in the discourse, as definites do, the distance between the last mention of the antecedent and the current NP form, that evokes the same representation, turns out to be crucial. This is usually a matter of recurrence, that is, later mentions are more accessible than earlier ones. The conclusions Ariel arrives at, concerning the contexts of occurrence of each expression, are that definite descriptions are found in the farthest contexts (same and across the paragraph boundary).

To account for the levels of accessibility, Ariel (1988, 1991) refers not only to the distance of the antecedent, but argues that universally, in the process of linguistic coding of the cognitive concept “degree of accessibility”, three criteria are mainly used: informativity (amount of lexical information), rigidity (ability to pick out a single referent, based on form) and attenuation (phonological size). The more informative the form, the more rigid and the less attenuated, the lower the accessibility, and vice versa.

With respect to the first criterion (informativity), the more lexical information the marker provides, the more suitable it is for retrieval of less accessible material. The more semantically empty the marker, the higher the level of accessibility of the referent. Definite descriptions have a marker with no descriptive content, which is followed by a full description. This confirms the different degree of accessibility between long and short definite descriptions (e.g., ‘*The Israeli linguist’* vs ‘*The linguist’)*.

With respect to the criterion of rigidity, it should be said that this criterion depends on the context, but nevertheless some markers are more “rigid” than others. Thus, so-called unique reference expressions can be used to retrieve entities that have a relatively lower degree of accessibility. This criterion distinguishes proper names from definite descriptions in the hierarchy (e.g., ‘*Ellen Prince*’ vs. ‘*The linguist*’).

The last criterion refers to form: the longer, phonologically heavy it is, the less accessible its referent will be. This criterion correlates closely with informativity, as we tend to be wordier to convey more information. These two criteria combined further confirm that definite expressions are

not a homogeneous category with respect to accessibility; the fact that they can retrieve referents at different distances in the text (Givón, 1992) and with different degrees of accessibility should be explained by referring to several other factors: topicality/relevance of the referent, syntactic role, status of the antecedent, etc. Ariel (1996) goes so far as to distinguish definite of 1, 2, or 3 content word, and definite NP + relative, which can also be used in first mention, since the new referent is very inaccessible and needs quite a lot of lexical material to be identifiable (as already noted by Chafe, 1976).[7]

Although the behavior of the forms varies from language to language, what remains constant and cross-linguistically universal is their association with the degree of accessibility and the hierarchy. Epstein (2002) builds on Ariel's work to conclude that definite articles behave more as markers of low accessibility rather than marker of givenness or identifiability.

Another proposal that tries to explain referential expressions in natural discourse is the one of Gundel et alii (1993), who proposes six cognitive statuses connected implicationally in a hierarchy, called the Givenness Hierarchy. The hierarchy of NP types is based on what speakers can expect about the cognitive status of the recipient of the NP denotation.

(9)

			uniquely		**type**
in focus >	activated >	familiar >	**identifiable** >	referential >	**identifiable**
{it}	{that this this N}	{that N}	{the N}	{indefinite this N}	{a N}

Givenness Hierarchy (Gundel et al., 1993:275)

[7] As for definite expressions, this criterion leads to unfulfilled expectations: in fact, definite expressions are more attenuated than demonstratives and thus should mark higher accessibility, when this is not the case. Ariel (2001) explains the attenuation of definite expressions by referring to their high frequency and the lower frequency of demonstratives.

The hierarchy in (9) is intended to linearly express implicational relations: each category imposes all the constraints of all the categories to its right, plus something else. Again, we will not analyze in detail each cognitive status presented, but we will focus on those that correspond to the (in)definite expressions, namely “type identifiable” and “uniquely identifiable”.

The cognitive status labelled as "type identifiable" is characterized as follows: the addressee is able to access a representation of the type of object described by the expression. This status is necessary for the appropriate use of any nominal expression and is sufficient for the use of the indefinite article ‘*a*’ in English. Thus, ‘*a dog’* in (10) is appropriate only if it can be assumed that the addressee knows the meaning of the word ‘*dog’* and can therefore understand what kind of object the phrase ‘*a dog’* describes.

(10) *I couldn't sleep last night**. A dog** (next door) kept me awake.*

The fact that the authors characterize cognitive status in terms of necessity and sufficiency is connected to the degree of freedom that the speaker can exert in the choice of referential expressions (as already noted by Du Bois, 1980). In this sense, since the status “type identifiable” is necessary for all nominal expressions, it can be coded in various ways, even different from indefinite descriptions. In languages such as Spanish, this status also corresponds to bare plurals or even to definite descriptions when it is not the referent as an individual but as a type to be identifiable (see1.3).
The cognitive status labelled as “uniquely identifiable”, on the other hand, is characterized as follows: the addressee is able to identify the referent on the basis of the nominal alone. This state is a necessary condition for any definite reference and is both necessary and sufficient for the appropriate use of the article ‘*the’* in English.

(11) *I couldn't sleep last night. **The dog** (next door) kept me awake.*

Identifiability may be based on a representation that already exists in the memory of the addressee, as would probably be the case in (11) without the material in brackets, but the authors (1993, 2001) point out, along with Hawkins (1978), that identifiability does not need to be based on prior familiarity, if a sufficient amount of descriptive content is encoded in the nominal itself. This is in line with the predictions made by Ariel's criterion of informativity.

For example, the phrase '*the dog next door*' in (11) would be perfectly felicitous even if the addressee did not previously know that the speaker's neighbor has a dog. Gundel et alii (1993) argue that with referential and uniquely identifiable NPs the hearer is able to retrieve the appropriate mental representation, based only on the referential expression itself. In contrast, with those NPs that are not uniquely identifiable, the hearer must rely on the content of the expression along with the rest of the sentence.

To conclude, what stems from these proposals is that definite expressions arise from the interaction between the participants in the discourse and help the mental processing of information. The processing is guided by pragmatic principles. As we see, these are ultimately linked to Grice's maxim of quantity "the speaker/writer is being as informative as required" (Grice, 1975).

2.3 The importance of definiteness in narrative discourse

Many of the studies we have mentioned here (Ariel, 1988; Chafe, 1976; Du Bois, 1980; Gundel et al., 1993) have been conducted on corpora of narrative discourse (spoken or written). Indeed, definiteness plays a crucial role in the construction of narrative discourse, being a category over which the speaker can exert a certain degree of control, and which guarantees communicative success. Definite expressions help to establish clear references, maintain textual coherence, and guide the hearer in understanding the characters and events in the story. In the narrative various characters, with varying degrees of importance, objects and actions are introduced, and definiteness helps the interlocutor identify the entities being referred to and keep track of their identity throughout the discourse. Definiteness serves not only to make sure that the addressee does not get lost in the

unfolding discourse, but also to draw his or her attention to particularly salient elements. As already seen, definite coding also serves to create or recall shared knowledge among participants. Consistent and appropriate use of definite encodings thus contributes to the overall cohesion of the text, which is well organized and more easily followed. It is a crucial mean in shaping the overall narrative experience.

3. Definites: core function and degrees of identifiability

As mentioned at the outset, the correlation between the cognitive category of identifiability and the grammatical category of definiteness is not perfect, that is, there is no 1:1 correlation between the identifiable or non-identifiable status of the referent and the definite or indefinite encoding of the NP designating that referent.

The use of the definite or indefinite article varies greatly from language to language. Languages that possess definite and indefinite articles often present a three-way distinction: definite article, indefinite article, and bare nominal, but the reference types under which these three options are allowed are not the same in all languages (Lambrecht, 1994).

However, it can be argued with Du Bois (1980) that the core function of articles is to mark an identifiability contrast and is restricted to those cases where the contrast is semantically possible. Identifiability is the core function of articles and is not applicable to non-referential noun phrases. Diachronically, in languages such as English or Italian, the article has gradually expanded its uses, including to contexts where the contrast of identifiability is no longer applicable, losing its primary function and becoming neutralized through predictability or interchangeability. We define these uses of the article as progressively more peripheral, in the sense of being away from the core function or less prototypical: starting with referents with partial identifiability, moving through uses that have more to do with content than with packaging such as generics (Chafe, 1976), up to nonreferential uses, in which the article no longer serves its basic function but becomes a nominal marking (Du Bois, 1980). Lyons (1999) too suggests that in some languages, such as English, we observe an

extension of the definite article to non- fully identifiable contexts, which are non-prototypical uses. Identifiability remains at core, but further uses of the definite article that are unrelated to identifiability are to be expected.

The distinction between grammatical definiteness and cognitive identifiability makes it possible to distinguish between a discrete and a non-discrete category. The definite-indefinite contrast is binary, that is, it is a matter of yes or no, whereas identifiability is a matter of continuum, in which different degrees can be identified.

A referent can be considered more or less identifiable, depending on various factors, especially psychological; while an NP cannot be more or less definite (Lambrecht, 1994). Each language, therefore, establishes different cut-off points for the lexicalization of articles (Dayal, 2004) along the identifiability continuum. Although grammatical encodings tend to be binary, formal boundaries between definite and indefinite arise from a psychological need for grammatical expressions of intermediate degrees of identifiability. It is quite common for referents with partial identifiability to be encoded as definite, just as those with total identifiability.

In the next section we will see the reasons why referents that are not fully identifiable are generally marked with the definite article, presenting one of the solutions proposed in the literature to the mismatch between form and function.

3.1 Uses of definite: sources of identifiability

In Givón's (2017) discussion of the mental processing sequence (see Section 1.2.1), we saw that one of the steps for encoding definites was to find the source of identifiability. The speaker may assume that a referent is identifiable to the hearer on the basis of both the external and internal context of the text and the general knowledge he/she possesses (that is culturally determined). This three-source classification, though sharable and cognitively based, needs to be expanded and more finely graded.

Here we will follow the proposal by Lyons (1999), which goes back to Hawkins (1978)[8]. The authors discuss four types of use of definites: situational, based on general knowledge, anaphoric, and associative.

Consider the following examples (Lyons, 1999:3):

(12) *Just give **the shelf** a quick wipe, will you, before I put this vase on it.*

(13) *I hear **the prime minister** behaved outrageously again today*.

Let us imagine that the sentence (12) is uttered in a room with a shelf. The shelf is immediately and unambiguously detectable by the hearer because it is present in the context of utterance. That is, the source of definiteness is the immediate situation, i.e., the extralinguistic context.

In the example (13) the relevant situation is larger: it is the whole country. When the term '*prime minister*' is used in a certain country, the prime minister of that country is typically meant. The person in question is not directly known to the hearer but is still identifiable, because it is reasonable to assume that every citizen knows that his/her country has a prime minister and knows who he/she is. Therefore, it can also be said that identifiability stems from the knowledge of the speaker. Consider now the following example (Lyons, 1999:3):

(14) ***The moon** was very bright last night.*

This example can be seen as a case of a broader situation, where the context is the whole world or the universe, but also as one of identifiability based on general knowledge. The referent is immediately identifiable because it is part of everyone's general world knowledge that there is one and only one moon in the sky, which the speaker is referring to.

[8] For a different classification, see Becker (2018) among others.

The three examples already show a continuum along which the situational context grows wider and wider.[9] Look at the following example (Lyons, 1999:3):

(15) *An elegant, dark-haired woman, a well-dressed man with dark glasses, and two children entered the compartment. I immediately recognized* ***the woman****.* ***The children*** *also looked vaguely familiar.*

In (15) the referents were first introduced at the beginning of the discourse and, since they are new information, they have been marked as indefinite. When they are reintroduced for the second time, they are marked as definite. This is the case of anaphoric use, in which the referent is identifiable because it has already been established; thus, the source of identifiability is internal to the text, i.e., it is the previous discourse. If we look at the following example, we notice a type of source of identifiability that, on the continuum, lies between anaphoric use and general knowledge:

(16) *I had to get a taxi from the station. On the way* ***the driver*** *told me there was a bus strike.*

Sentence (16) is an example of associative usage, also called bridging cross-reference or simply bridging. The driver is unknown to the listener; in fact, if we suppose to put him in front of a series of taxi drivers' faces, he would not be able to identify which one was driving the vehicle that his interlocutor has taken. It should be stressed, however, that it is part of everyone's general knowledge that each taxi is always associated with one and only one driver. The use of definite presented in sentence (16) is a combination of anaphoric use and use based on general knowledge. It is precisely the ability to build this association, based on the hearer's encyclopedic knowledge, between the referent previously introduced (*taxi*) and the new referent (*driver*), that is the source of this kind of

[9] Recall what was said in section 1.2.1 about the cognitive correlates of retrieving these referents, which are stored and retrieved in different ways.

definite status. The driver is identifiable in the sense that he/she is the driver of the particular taxi under discussion. The same explanation is appliable to the following cases in Spanish[10]:

(17) <u>*Me*</u> *duel-e* ***la*** ***pierna*** (Leonetti, 2019:4)
me.DAT. hurt.PRS.3SG the leg

(18) <u>*Pedro*</u> *perdió* ***el*** ***móbil*** *esta* *mañana*. (Espinal & Cyrino, 2017:2)
Pedro lost.PST.3SG the cell phone this morning

In the examples above, the definite referent (in bold) is associated with the previous referent (underlined) through a special kind of bridging, which is the inalienable possession relationship[11]: body part in (17) and personal possession in (18) (Givón, 2017; Schwarz, 2013).

Now consider the following example:

(19) *I've just been to a wedding.* ***The bride*** *wore blue*. (Lyons, 1999:7)

the sentence can be considered an example of bridging cross-reference since the definite status of '*the bride*' is licensed by its association with a previously introduced referent (*the wedding*).

Lyons (1999) brings up this example to argue that it is problematic to associate this kind of use of definite NP with the identifiability of the referent by the hearer. "Is it accurate to say that the hearer identifies the referent in some real sense? He doesn't know who she is yet and he doesn't know anything about her. If he were asked later who got married that morning he would not be able to tell based on [the example 19], and if he passed the bride on the street the next day he would not recognize her as the person referred to." (Lyons, 1999:7).

[10] The abbreviations comply with the Leipzig Glossing Rules, available at https://www.eva.mpg.de/lingua/resources/glossing-rules.php (see Abbreviations).
[11] This type of definite has been widely discussed in the literature and often goes under the label of expletive definite. See Brugger (1993); Espinal & Cyrino (2017); Le Bruyn (2014); Longobardi (1994); Vergnaud & Zubizarretta (1992).

Other authors argue that examples such as (19) can still be linked to the notion of identifiability. Evers (2020) discusses the same example and argues that bridging reference is both identifiable and definite and that identifiability is still the core information. In fact, what is salient in the discourse is the role of bride, not her personal identity, and this is well known as part of the hearer's general knowledge. In the hearer's prototypical mental representation of a wedding event, the bride is definitely an identifiable component. The use of a definite marking suggests that hearers, regardless of their ability to recognize the bride on the street, can rely on their default mental representation of the event to identify the referent in the context of the discourse. The use of an indefinite article '*a bride*' would have instead implied that the person in question was not prototypically associated with a wedding and that perhaps there was more than one bride at the event, wearing different colors. It is the reference to an identifiable component of the hearer's internal representation of an event that makes appropriate to mark such a referent as definite.

What has been called "internal representation of an event" here is generally referred to in the literature as cognitive schema or frame. Fillmore's (1982:111) definition is provided below:
"By the term "frame" I have in mind any system of concepts related in such a way that to understand any of them you have to understand the whole structure in which it fits; when one of the things in such a structure is introduced into a text, or into a conversation, all of the others are automatically made available."

The lack of full identification for referents which are encoded as definite, is compensated by deriving their identifiability from semantic frames; no further information is needed[12].

It is clear from this example that there are different degrees of identifiability, which are arranged along a continuum, from the degree of fully identifiable to the degree of non-identifiable, with boundaries that are not always clearly demarcated. Speakers of languages in which identifiability

[12] Ariel (2001:33) points out that within the same frame, not all components have the same degree of accessibility; those less salient or necessary for the specific frame are less accessible.

is grammaticalized in terms of definiteness generally have to decide whether to encode entities with variable degrees of identifiability in definite or indefinite forms. The boundary between definite and indefinite encoding on the continuum of identifiability is not always obvious in any language. Like a referent of full identification, a referent of partial identification is usually designated as identifiable and given a definite encoding. It is noteworthy that, in instances of partial identification, in languages lacking an explicit definiteness the referent takes the form of an indetermined expression (Chen, 2009). We will see later (2.4) the case of Mandarin Chinese.

By referring to the continuum, cases in which the referent is not fully identifiable by the hearer but are nonetheless marked with a definite article become explicable. See the examples discussed below:

(20) *The boy scribbled on* ***the living-room wall.*** (Du Bois, 1980:232)

(21) *I am going to* ***the cinema***. (Leonetti, 2019:3)

(22) *I met* ***the daughter*** *of a king.* (Chen, 2004: 1142)

(23) ***The mayor*** *of a small village in Wales.* (Chen, 2004:1141)

(24) *The baby grasped* ***the finger*** *of the surgeon*. (Espinal & Cyrino, 2017:2)

(25) *The driver lost the control of the vehicle when* ***the wheel*** *of the truck exploded*. (Espinal & Cyrino, 2017:2).

According to Du Bois (1980), in (20) the ability of the addressee to precisely determine which of the living room's four walls is involved is not needed for the definite encoding of the referent of '*living-room wall*'. The hearer can identify the precise living room in question and narrow the list of probable referents to just one of the room's four walls (without knowing exactly which), and this will be sufficient for the communicative exchange at hand. Moreover, an addition of information would lead to the violation of Grice's maxim of quantity, which would be equally violated if the referent is coded

as non-identifiable. This is consistent with the pragmatic principles guiding the choice of referent in discourse and their mental processing proposed by Ariel and Gundel (see. 1.2.2).

To explain cases like (20), Du Bois (1980) proposes the curiosity principle, which is formulated as follows: "a reference is counted as identifiable if it identifies an object close enough to satisfy the curiosity of the hearer. In special circumstances even an ordinary speaker might desire more precise identification. But in everyday speech such partial identification is quite common." (Du Bois, 1980: 233).

If we consider the examples (21) - (25), we notice a mismatch between form and function: the mismatch is the impossibility to satisfy the conditions of identifiability, encoded by the definite article at the level of token referents (Leonetti, 2019). Those are some examples of weak definites[13]: a short weak definite is in (21), long weak definites are in (22) - (25) (with (24) and (25) representing cases of part-whole relationship).

Again, the resolution involves satisfying the identification at the more abstract level of types or roles/functions in a frame (Leonetti, 2019:23). Frame-triggered referents can stand in a one-to-one connection to the anchor, like a driver to a car or a prime minister to a country (see example (5) where the whole country and its governmental set can be considered the activated frame; what Hawkins calls "wide situation"), or in a one-to-many relationship, like daughter to a father or a finger to a hand (Löbner, 1985).

What matters here is not the identification of the particular cinema or the finger of the hand, nor providing information on the number of king's daughters, what matters to the speaker is his interlocutor being able to identify that type of referent and understand its role within the activated frame. Thus, the success of the intended reference depends on the hearer's cooperation (Lyons, 1999).

[13] In the literature weak definites are usually divided into two groups: short and long weak definites. The first type contains only the definite article and a common noun (es. read the newspaper, play the violin); in the second group the NP has follows more complex scheme: "the N of {a/the} N" (es. the picture of a young couple). For a description and discussion also from a crosslinguistic point of view, see Aguilar-Guevara et alii (2019); Leonetti (2019).

4. Formal means of expressing definiteness

This section first describes how languages express definiteness and then more attention will be paid to the languages involved in the study. As already stated, we can safely assume that all languages must be capable of expressing definite reference and that, as a result, there must be a way in every language to refer to particular individuals who are assumed to be known to speaker and hearer in the relevant context of a speech act (Aguilar-Guevara et al., 2019). In this section an attempt will be made to give a brief overview of the main formal means that languages use to mark a referent as definite or indefinite.

Definiteness is one of the domains in which linguistic diversity most clearly manifests itself, as the languages exhibit very different solutions. Given this wide variability in formal means, it is useful here to take up the distinction proposed by Lyons (1999) between simple and complex (in)definiteness. In many languages the NP contains an element that seems to have the main function of encoding (in)definiteness: the article. It is true, however, that many NPs that do not contain an article still give the possibility of a definite or indefinite interpretation: e.g., in '*this closet*' the referent is identifiable, while in '*some closet*' it is not. What Lyons (1999) argues is that the NP with article can be considered the main instantiation of the definite NP, since definiteness stems precisely from the article itself. This type of NP will be referred to when we speak of simple definite. By complex definite, on the other hand, we refer to NPs with items, other than the article, that primarily encode other semantic-pragmatic categories (e.g., deixis, possession, quantification etc.), but from which in any case the interlocutor understands whether or not he/she is able to identify the referent.

Concentrating on simple definites, it is possible to make a first, general division among the languages of the world: languages that have some form of article and languages that have none. Among those that have articles, one can identify languages that have only the definite article, languages that have only the indefinite article, and languages that have both (Lyons, 1999).

4.1 Simple Definites

Dryer (2013) identifies two types of languages, among those having a definite article:

- Languages in which the definite article is a word distinct from the demonstrative
- Languages in which the definite article is an affix on the noun

The first distinction we can make is thus morphological, between articles that are free morphemes and articles that are bound morphemes.

Within the family of Romance languages, most belong to the second type, while Romanian belongs to the first. See the examples from Italian (26) and Romanian (27) for '*the man'* (Giusti, 1994: 241).

(26)	***l'***	*uomo*
	the.M.SG	man

(27)	*om-*	***ul***
	man-	the.M.SG

Another variable often considered when describing article systems is the position relative to the noun: preposed or postponed. This trait intersects with the one presented earlier. Romanian is an example of a postponed bound article (27), while in Modern Standard Arabic the definite article is a preposed affix, see the example below for '*the house'* (Lyons, 1999:50):

(27)	***al-***	*-baytu*
	the-	-house

If we compare the Romance article with the Arabic or English article we realize in the former case that in addition to identifiability other grammatical categories, such as gender and number, are encoded, while in the latter cases the morphemes are neutral with respect to any other grammatical category. Thus, we can have languages in which the article is an inflected form, whether it is a free morpheme or an affix, and languages in which it is an invariable morpheme, either free or affix.

Although they are frequently able to be stressed for emphasis or contrast, definite articles exhibit a strong propensity to be unstressed. This reality manifests itself in the fact that the article is a weak form in many languages that use phonological reduction strategies. Additionally, articles are frequently monosyllabic, due to the strong tendency for unstressed words to be monosyllabic and the correlation between stress reduction and the tendency of polysyllabic items to be reduced to monosyllabicity (Selkirk, 1984). This obviously affects the perceptual salience of these items in the speech stream[14].

4.2 Simple indefinites

For the indefinites Dryer (2013) identifies three types of languages:

1. Languages with an indefinite article distinct from the numeral for 'one'[15]
2. Languages with an indefinite article identical to the numeral for 'one'
3. Indefinite article affixed to the noun

As with definites, we distinguish the first two types in which the articles are free morphemes, as in the case of English (28), belonging to type 1, and German (29), belonging to type 2 (examples from Dryer, 2013), and the last type in which the article is a bound morpheme, as in the case of Modern Standard Arabic (30) (from Lyons, 1999:50):

(28)	***a***	*woman*	(29)	**einen**	Hund	(30)	*bayt-u-*	***n***
				a.M.SG.ACC	dog		house.NOM-	a

[14] However, since it depends on the language's suprasegmental phonological structure, the tendency for definite articles to be unstressed is not universal. Only languages where there is word stress can have a distinction between fully stressed "lexical" words and frequently unstressed "function" words (Lyons, 1999).

[15] It is therefore not easy to distinguish when the form is an indefinite or a cardinal expression; in these cases, very often the obligatory nature of the indefinite in certain semantic contexts, in the absence of other determinants, helps the disambiguation (Lyons, 1999:95)

Other important traits are inflection and position, which, can be prenominal or postnominal. In the case of English, the indefinite article is not inflected, while in the case of German it is sensitive to the categories of gender, number, and case. As the typology just presented shows, there are languages in which the numeral for 'one' and the indefinite article are the same, at least phonologically.

It is useful here to note that there are languages, such as French and Italian, that have a series of articles called partitive articles. In these languages the singular countable noun takes an indefinite article identical to the numeral, while with mass nouns and indefinite plurals the partitive article is used. Italian also allows bare indefinite plurals and masses, while French generally does not (Belletti, 1988)[16].

4.3 Languages without articles

As already stated, we should expect all languages have some means of expressing definiteness, even without overt dedicated elements. Again, the level of variation is very high, with different kind of means:

- **Phonetic**: an example is Czech, which uses intonation to mark a referent as identifiable (Cummins, 1998).

- **Morphological**: Comrie (1978) describes some languages in which there is a prepositional or postpositional object marker, recurring only with definite object NPs (this is the case with Hebrew). Other languages use an agreement marker outside the NP, such as object-verb agreement in Uralic languages. Hungarian has two paradigms for each transitive verb: one called definite and one indefinite (Bánhidi et al., 1965).

- **Syntactic**: examples are the use of a pronominal marking for definiteness in Yoruba (Rowlands, 1969) and word order strategy. The latter is particularly prevalent. Examples include Czech, Chinese,

[16] It should be noted that bare nominals are allowed in many languages. In languages without articles, it is obviously more complex to interpret bare NPs. For a comprehensive analysis in this respect see Dayal (2004; 2018), Despic (2019).

and Cantonese (Cheng & Sysbema, 1999). In Czech bare nominals and cardinal NPs in sentence-initial position receive a definite interpretation, while the default reading in final position is indefinite.

4.4 Complex definites

At the end of this section complex definites are presented, although a detailed analysis is beyond the scope of this chapter. We just provide an informal definition and a list of possible realizations[17].

Complex definites are said to be those nominal phrases whose definite interpretation is due to inherent definiteness or something other than the article, be this another kind of determiner or a modifier. Complex definites are (as listed in Lyons, 1999):

- proper nouns
- personal pronouns
- demonstratives
- possessives
- universal quantifiers

Here we only provide a brief description of two phenomena: the relationship between definite article and demonstratives and the linguistic typology of possessives.

As for demonstratives, in some languages they are in complementary distribution with definite articles, while in other languages these two kinds of determiners can also cooccur (Guardiano, 2012). This is the case in Lakhota, see example (31) from Ingham (2001:16) (the article is in bold, and the demonstrative is underlined):

(31) wicʻaṡa **ki** <u>he</u>

man the that

[17] Complex indefinites will not be discussed in this chapter. See Gianollo & Silvestri (2022: 695-727); Haspelmath (1997); Lyons (1999:148-152).

Dryer (2013) also includes in his typology languages in which demonstratives are used as a marker of definiteness. Here we refer mainly to the fact that in these languages (e.g., Swahili and Tigrinya to name a few) the demonstrative is used with anaphoric function, with very high frequency, more so than in English. The use of demonstratives together with an NP with an anaphoric function is quite widespread in the world's languages, to the extent that it is possible to observe a continuum ranging from rare use of the anaphoric demonstrative to very frequent use. Using a demonstrative, where languages that possess it would use an article, is also a strategy used in many languages lacking any overt marking of definiteness (such as, for example, Chinese and Cantonese, see also Becker, 2021).

Our second point of interest is the linguistic typology of possessives as presented by Lyons (1999:24-26). The author divides the world's languages into two types:

1. Determiner-Genitive (DG) languages.
2. Adjectival-Genitive (AG) languages.

In the former, possessives appear in positions normally reserved for determiners and induce a definite reading in the NP. An indefinite reading is therefore impossible with such a basic structure. English and French are DG languages. To give an indefinite reading, that is, when the reference is not identifiable, a prepositional construction is generally used. See the following examples in English:

(32) ***my*** *friend* (definite reading)

(33) *a friend* ***of mine*** (indefinite reading)

In AG languages, such as Italian or modern Greek, the possessive is in adjectival (or similar) positions and may cooccur with the definite article for a definite reading or with the indefinite article for the indefinite interpretation. See examples (34-35) from Italian (for a discussion of determiners in Italian see 2.1). In conclusion, possessives can be considered definite determiners only in DG languages.

(34)	***la***	*mia*	amica		(35)	***una***	*mia*	*amica*
	the.F.SG	my.F.SG	friend.F.SG			a.F.SG	my.F.SG	friend.F.SG

2. Definiteness in the languages of the study

After the analysis of definiteness on the discourse-pragmatic level and an overview of the main formal means used by the languages of the world to encode it, in this chapter, a description of the formal strategies used in the family languages involved in the study: Italian, Spanish, Arabic is provided. In the research presented in the following chapters, the data that will be analyzed were collected in Italian, but since the experimental group is composed of children with different family languages, which could be a source of crosslinguistic influence, it seemed useful to give a description of them as well. Section 1 is dedicated to Italian, section 2 is dedicated to Spanish, section 3 to Arabic, specifically its spoken varieties, and section 4 to Chinese, concluding the chapter.

For each language, I will give a morphosyntactic description of the formal strategies used to express definiteness and the contexts in which the presence of the article (or other formal element) is obligatory or not. In addition to a distributional description, an attempt will be made to make explicit the interpretation that the nominal phrase may receive, if it is referential (otherwise the identifiability contrast does not apply).

1. Italian

Italian has both a definite article system and an indefinite article system. In both cases articles are prenominal free morphemes, sensitive to the categories of gender (masculine and feminine), number (singular and plural) and subject to allomorphy (Chini, 1995; see Table 1 in Appendix A). The occurrence of the forms of the masculine singular *il*/*lo*/*l'* (definite), *un*/*uno* (indefinite) and plural i/*gli* (definite) are conditioned by phonological factors: *l'* is needed with nouns beginning with a vowel, *lo* ,*gli* and *uno* with those beginning with the cluster [s] + consonant or the sounds [ʃ], [ks], [ts], [dz], [ɲ], [pn], [ps], while *il* and *un* are the basic forms for all other masculine nouns. *La*, *l'*, *le*, *una* and *un'* are forms of the feminine, which are also subject to phonological restrictions, so the elided forms

occur with nouns beginning with vowels. Italian, as already mentioned, has a system of partitive articles constructed with the *de + definite article* form, which in this case is a mere marker of gender and number (see Table 2 in Appendix A). Plural partitive articles are treated as plural forms of the indefinite article and are therefore expected in unidentifiable contexts (Cardinaletti & Giusti, 2016).[18]

Regarding obligatory contexts, Italian requires the presence of the determiner with singular countable nouns in the main argument positions (subject, direct object[19], prepositional object, inverted subject of either ergative or unergative predicates); see the following examples from Longobardi (1994:612):

(35) a. **(Un/Il)* ***grande amico*** *di Maria mi ha telefonato.*

(a/the) ***great friend*** *of Maria called me up.*

b. *Ho incontrato *(un/il)* ***grande amico*** *di Maria ieri.*

I met (a/the) ***great friend*** *of Maria yesterday.*

c. *Ho parlato con *(un/il)* ***grande amico*** *di Maria ieri.*

I spoke with (a/the) ***great friend*** *of Maria yesterday.*

d. *Ha telefonato/E' venuto *(un/il)* ***grande amico*** *di Maria.*

called up/came (a/the) ***great friend*** *of Maria.*

This type of noun cannot occur without a determiner, but this constraint does not work for nominals in typical nonargument function, as in vocative, predicative, or exclamatory contexts (Longobardi, 1994)[20]. Plural nouns in Italian can occur with the definite article (36a), with the partitive article (36b), or bare (36c). In the first case the article can signal either the identifiability of the referent or a

[18] For a different treatment, see Zamparelli (2008).

[19] Renzi (1988:376-377) points out that a singular countable nominal phrase can be determinerless when it is object, it is particularly heavy (i.e., consisting of multiple elements) and in preverbal position (this usage is marked high on the diaphasic axis and typical of written style).

[20] Modification often nullifies this restriction and requires the use of the determiner (Renzi, 1988:405-406).

generic reading (as will be seen later in 42), in the second and third case the reading is indefinite (Cardinaletti & Giusti, 2016, 2018).

(36) a. *Ho visto* ***i ragazzi.***

I have seen ***the boys****.*

b. *Ho visto* ***dei ragazzi.***

I have seen ***de.ARTDEF boys****.*

c. *Ho visto* ***ragazzi****.*

I have seen ***boys****.*

Similar is the behavior of mass nouns that can occur with the definite article (37a), with the partitive article (37b), or as bare nominals (37c). See the following examples from Cardinaletti and Giusti (2018):

(37) a. *Ho raccolto* ***il fieno.***

[I] have harvested ***hay.***

b. *Ho raccolto* ***del fieno***

[I] have harvested ***de.ARTDEF hay****.*

c. *Ho raccolto* ***fieno.***

[I] have harvested ***hay.***

According to the authors, with mass nouns "the definite article is ambiguous between definite and indefinite meaning" (Cardinaletti & Giusti, 2018:142), the partitive also in this case raises an indefinite reading and the same applies to bare mass nouns in argument position (postverbal if subject) (Cardinaletti & Giusti, 2018; Longobardi 1994). According to Cardinaletti & Giusti (2018), disambiguation of the interpretation of the definite article with mass nouns would come from the

context, particularly from the interaction with verbal aspect: the definite interpretation would be ungrammatical if the event is telic (see Cardinaletti & Giusti, 2018:142-146 for examples and a more in-depth discussion).

More difficult is to characterize the sometimes-alternating use or omission of the article in certain prepositional complements in Italian[21]. Lyons (1999:51) points out that bare nominals with prepositions are common in many languages and that explanations for these patterns remain mysterious. Giusti (1994) similarly argues that it is a general phenomenon (discussing data from Romanian and English as well as Italian) and that explanations for the absence of the article are often missing (also noted by Longobardi, 1994:612). In any case, Giusti states that "some prepositional complements in Italian must have no article when they embed a bare noun with a certain interpretation which […] we may label as salient in some sense":

(38) *Siamo in **giardino.***

*[We] are in **garden.***

The example (38) given by Giusti is part of a complex list of locative prepositional formations, presented by Renzi (1988:396, 412) who argues that the absence of the article, in the face of the possibility of its presence in the same type of phrase, reveals high familiarity of the referent (the referent is particularly accessible, in other terms). The character of these formations can be both abstract and concrete:

(39) *a. Si sono sposati **in comune*** (abstract) *b. Ci vediamo domani **in comune*** (concrete)

*[They] got married **in** (the) **town hall*** *See you tomorrow **at** (the) **town hall***

[21] Both with countable and mass nouns.

In (39a) the institution is intended, in (39b) the physical place. The contrasts of salience, familiarity and concreteness may also be useful in explaining PP pairs with and without determiner, especially with singular or mass nouns: '*in acqua/nell'acqua'* (in the water), '*in mare/nel mare'* (in the sea), '*in cielo/nel cielo'* (in the sky) (but more in-depth analysis is needed).

There is a third type of bare noun occurring in Italian in argument function[22]: some rarer cases of singular countable nouns in the scope of a sentential negation[23]; but these cases are considered quasi-idiomatic expressions for the purpose of this work and will not be considered in the following analysis. See example (40) from Longobardi (1994:613):

(40) *Non ha proferito* ***verbo****.*

he didn't utter (a) ***word.***

Idiomatic expressions, as we have already seen in 1.1.1, are part of nonreferential expressions and contain bare nominals, definite and indefinite NPs, even in Italian (Renzi, 1988:344)[24]:

(41) a. *fare* ***il bagno***
to do ***the bath***

b. *fare* ***un bagno***
to do ***a bath***

The generic reading in Italian always involves the use of the article, which can be definite (42a, b) or indefinite (42c, d) if the noun is countable, definite if it is mass (42e)[25] (Longobardi, 1994:614-631):

[22] A separate treatment should be reserved to nouns in coordination, which can occur bare or with determiner. For further discussion, see Heycock and Zamparelli (2000, 2003, 2005).
[23] For an examination of the restrictions of countable bare nouns in negative sentences see Renzi (1988:382-383).
[24] Renzi (1988:407-418) proposes a rich list of idiomatic expressions and a comprehensive examination of other cases in which the noun admits or disallows the article (with light verbs, with other prepositions, in adverbial phrase). These do not appear in our corpus, often are non-referential and will not be discussed in this section.
[25] See the discussion between Longobardi (1994) and Chierchia (1998) on the possibility of Italian bare nouns in argument position to receive generic reading.

(42) a. ***Il dinosauro*** *fu ucciso da cause misteriose*

the dinosaur *was killed by mysterious causes.*

b. ***I dinosauri*** *furono uccisi da cause misteriose.*

the dinosaurs *were killed by mysterious causes.*

c. ***Un cane grosso*** *crea sempre questi problemi.*

a large dog *always raises such problems.*

d. ***Dei cani grossi*** *creano sempre questi problemi.*

De.ARTDEF *large dogs always raise such problems.*

e. ***L'acqua*** *di quel colore raramente può essere bevuta.*

the water *of that color can rarely be drunk.*

As has already been mentioned, Italian requires the use of the article even with possessives, except for singular kinship terms: *mia mamma/mio papà/suo fratello (*my mum/my dad/his brother) (Renzi, 1988:398). Only the definite article can occur with numerals (43a); numeral + NP without article raises an indefinite reading (43b) (Renzi, 1988:366).

(43) *a.* ***I tre dottori*** *hanno visitato il paziente*

The three doctors *examined the patient*

b. ***Tre dottori*** *hanno visitato il paziente*

three doctors *examined the patient*

In standard Italian, the article does not cooccur with demonstratives (Calabrese, 1988: 618-619).

2. Spanish

Spanish, like Italian, has both a system of definite articles and indefinite articles. In both cases these are prenominal free morphemes, sensitive to the categories of gender (masculine and feminine) and number (singular and plural; see Table 3 in Appendix A). With feminine nouns beginning with a

stressed vowel, the feminine article takes the form *el*, instead of *la*: *(*la) el agua, (*la) el aguila* (the water, the eagle). Wherever Italian uses partitive articles, Spanish uses the plural forms of indefinites (Leonetti, 1999).

Spanish, like Italian, does not readily admit forms without determiners, which are normally licensed by a lexical head (V as in 44b or P as in 44c) (Chierchia, 1998). As in Italian, singular countable nouns always want the article when they are in the main argument positions. Bare countable singular nouns are therefore allowed in the following three cases (examples are from Lyons, 1999:104-105):

- the noun is not in one of the main argument positions (44a)
- the noun is governed by a series of verbal heads (44b)
- with some prepositions (44c)

(44) a. *Maria se hizo* ***dentista***

Maria became (a) ***dentist***

b. *Ana tiene* ***coche***

Ana has (a) ***car***

c. *Una casa con* ***jardín***

a house with (a) ***garden***

Verbs that admit bare nominals as objects are characterized by Espinal (2010) as atelic and listed by Butt and Benjamin (1994:46) under the label of "verbs that denote things of which one normally have one at a time": having, seeking, producing, etc. Prepositions that admit bare countable nouns are not easy to characterize, e.g., *sin*, *por*, *como* (without, by way of, as)[26]. In all cases, both Lyons (1999) and Espinal (2010) assert that the interpretation of these NPs is indefinite.[27]

[26] Compare the lists of prepositions provided by Butt and Benjamis (1994) for Spanish and Renzi (1988) for Italian to note converging patterns.

[27] Espinal (2010) addresses this issue extensively and argues that this type of bare nouns is closer to bare plurals than to singular indefinites with the article (both semantically and syntactically).

Plural countable nouns can be definite, indefinite, and bare. When they occur with the definite article, they can have a definite (45a from Butt & Benjamin, 1994:33) or generic interpretation (as will be seen later), when occurring with indefinites they have an indefinite reading (45b from Butt & Benjamin, 1994:49), while the plural bare nouns are generally indefinite (45c from Espinal, 2010:988):

(45) *a. Traeme **los tenedores** y **las cucharas*** *b. Le dieron **unas monetas*** *c. necesitar **zapatos**.*

*Bring me **the forks** and **the spoons.*** *[They] gave him **some coins*** *to need **shoes**.*

Bare plurals, in fact, denote properties or rather types (Chierchia, 1998; Espinal, 2010), corresponding to the cognitive status of "type identifiable" in the hierarchy of Gundel et alii (1993:284). Mass nouns behave as in Italian and may occur with the definite article (46a) for a definite reading or bare (46b) for an indefinite reading (Gerards & Stark 2022). See examples from Gerards and Stark (2022:4):

(46) a. *Derretir **la mantequilla*** fría b. *derretir **mantequilla** fría*

*to melt **the** cold **butter*** *to melt cold **butter***

Generic reading in Spanish always involves the use of the article, which can be definite or indefinite if the noun is singular countable, definite if it is mass or plural countable (Leonetti,1999:870-882). Regarding possession, Spanish is a mixed-system language, combining forms of DG and AG languages (Lyons, 1999:133). Spanish has short forms of possessive determiners that induce a definite reading in the noun and cannot appear with the article (47a); and long forms that can occur both with the definite (47b) and indefinite article (47c). Thus, to express indefinite possession, Spanish

normally resorts to a construction with the noun preceded by the indefinite article and followed by the long form of the possessive (Butt & Benjamin, 1994:97)[28].

(47) a. *En* ***mì novela*** b. *En* ***la novela mía*** c. *En* ***una novela mía***

In ***my novel*** *In* ***the novel (of) mine*** *In* ***a novel (of) mine***

In Spanish, it is possible for the definite article to cooccur with the demonstrative, when the latter is in postnominal position. The interpretation proposed by Leonetti (1999:811) is that such a construction signals a lower degree of accessibility of the referent, which therefore needs to be overspecified, and that stylistically it is used to indicate psychological distance between speaker and referent, see the example below:

(48) ***La*** *casa esta*

the *house that*

This construction is pragmatically quite marked and strongly grounded in specific communicative situations, where it serves to evoke a despective interpretation (although Brugé, 1996 asserts that this interpretation may not hold true for all speakers and in all situations).

We want also to note the case of nominal ellipsis, which is one of the features that differentiates the Spanish definite article from the Italian one. In Spanish the article can combine with adjectival phrase, e.g., '*el más caro*' (the most expensive), prepositional phrase, e.g., '*el coche, el de Teresa*' (the car, the one of Teresa) or with relative clause, e.g., '*la que quieres'* (the.F.SG. one you want). In many Romance languages these constructions are not possible, and Italian requires in most

[28] Lyons (1999:25) points out that in Spanish there is a clear distinction between the predicative form of the possessive and the definite reading of the long possessive: *Esa pluma es* ***mía*** (predicative non-pronominal possessive) vs. *Esa pluma es* ***la mía*** (pronominal possessive), corresponding to English *this pen is mine*.

of these cases the use of the distal demonstrative '*quel*', e.g., '*il/quello più costoso'*, '*la macchina, quella di Teresa'*, '*quella che vuoi*' [29] (Leonetti, 1999:818).

It is worth pausing briefly on the definiteness restriction that distinguishes two existential constructions in Spanish: the one with *haber* (49a) and the one with *estar* (49b). The former does not admit definite NPs, as often found in existential constructions in many languages, since it serves to introduce new referents into the discourse; the latter prefers definite referents and rarely admits indefinite ones (only if they have specific interpretation; that is why with *estar* bare plurals are not admitted, since they always receive existential interpretation) (Leonetti, 1999: 814-817 for a more detailed analysis of definiteness restrictions).

(49) *a. Hay (*està)* ***un error*** *en esta página*
there is ***a mistake*** *in this page*

*b. Aquí está (*hay)* ***el regalo*** *de Juan*
here (it) is ***the present*** *of Juan*

As in Italian, also in Spanish, the specific reading of NP with numeral wants the definite article, the absence of which brings out nonspecific reading[30] (Leonetti, 1999).

3. Arabic

As will be seen in Chapter 4, the families of the children involved in the study are from four Arabic-speaking countries: Morocco, Tunisia, Egypt, and Syria, which correspond to two macro-varieties of Arabic: Eastern and Western (Cotter & de Jong, 2019). What we are interested in here, therefore, is to provide a brief description of the spoken varieties of Arabic, not Modern Standard Arabic (MSA from here on), which does not constitute the language which the participants are exposed to. All four varieties of Arabic of our interest have a definite article (cf. example 50 from Brustad, 2000:19),

[29] It should be mentioned that Spanish knows another article, '*lo*,' often used in combination with the same type of AP or PP and relative clauses. The debate about this element is great and its semantic and syntactic features are complex, so we refer to Leonetti (1999:829-834) for a full description.
[30] For the use of the plural indefinite article with numerals, as an indeterminator, see Leonetti (1999:843-844).

which is bound and prenominal, (*al-* in MSA, (*i-*)*l* in spoken varieties) and which can undergo a process of assimilation of the coronal trait with the word to which it is bound.

(50) *fi* ***l-****bet*

in ***the****-house*

The distributional pattern of the definite marking is consistent across all varieties of spoken Arabic considered in this study (Turner, 2021). The definite article is allowed with both countable and mass nouns. Abstract nouns require the definite article as well (Brustad, 2000). Only a few semantically connoted categories of nouns do not generally occur with the definite article, among others: some kinship names, body parts, personal possessions (Harrell, 2004). These nouns establish an inalienable possession relationship with the possessor, which induces an inherent definite reading (Bustrad, 2000: 40).

(51) *'eyyett 'la* ***bba.*** (Moroccan)

called-I on ***father****.*

I called my father.

The central function of the article is to mark identifiability of the referent. However, the definite article in Arabic has developed other grammatical functions unrelated to its canonical use. It is optionally attached to nouns that refer to partially identifiable referents, it is used in environments where the opposition definite/indefinite is irrelevant (Jaber et al., 2022; Kashgary, 2015) and with non-referential nouns (Brustad, 2000)[31].

[31] According to Brustad, the use of the definite article in Moroccan would deserve a separate discussion, which cannot be addressed here. We limit our analysis to the features common to all spoken varieties. For a discussion of Moroccan see Brustad (2000:36-42).

(52) *šra magana dyal* ***l-ma.*** (Moroccan; Brustad, 2000:81)

bought-he watch GEN ***the-water.***

he bought a ***water****proof watch.*

(53) *fi wahid badwi, fat 'a* ***l-mat'am****.* (Syrian; Brustad, 2000:20)

there-is one bedouin, entered-he into the-restaurant.

there's a bedouin who went into ***the restaurant****.*

In example (52) '*l-ma*' (the water) is translated into English with a compound (waterproof watch), which in Chen's (2009) classification is counted among the non-referential qualitative NP forms (see 1.1.1). In Italian a literal translation could be '*orologio* ***per acqua*****'** (watch **for water**) with a bare nominal or less literally '*orologio resistente al****l'acqua***' (**water** resistant watch) with a definite (recalling the oscillation, mentioned above, between bare nouns and definite determiners typical of mass nouns with preposition). On the contrary, it is clear from the example that, even in these cases, Arabic requires the definite article. (53) is very similar to example (21) in paragraph 1.3.1, where '*cinema*' is an example of a partially identifiable entity, since what needs to be identified is the type and not the actual referent, exactly like '*restaurant'* in the example above. In this case, Arabic, Italian and English behave the same way (preferring a definite form). Conversely, in contexts in which English and Spanish would use bare plurals, that is, when referring to categories as a whole (Harrel, 2004:190), Arabic and Italian are aligned in requiring the use of the definite article (54):

(54) *sir-le-s-suq u sri* ***l-ful*** (Moroccan)

go to the market and buy (the) ***beans***

vai al mercato e compra ***i fagioli*** (Italian)[32]

[32] It should be noted that in Italian in such a context, bare plurals are not ungrammatical, although not preferred.

Even under the scope of negation, Arabic wants a definite, whereas in this case Italian prefers bare plurals. See the following example from Brustad (2000:36):

(55) *hada wahed ar-razel ma'anduš* ***l-wlad***. (Moroccan)
this one the-man neg at-him ***the-children.***
questo è un uomo che non aveva ***figli****.* (italiano)
(*this is a man who has no children)*

For generic readings, Arabic mainly uses the definite article (Jaber 2014; Kashgary, 2015).[33]

More complex is the expression of indefinite, which is traditionally said to be encoded by the suffix '*-n'* (known as "nunation" in Arabic studies), which is in semi-complementary distribution with the definite article and is sensitive to the category of case (with a three-case distinction: nominative, accusative, oblique) (Lyons, 1999). The debate on the status of this suffix is complex, especially on whether it can be considered marking true indefiniteness[34]. See the following examples in MSA (Lyons, 1999:50):

(56) a. ***al****-baytu*
ART.DEF-house
the *house*

b. *bayt-u-****n***
house-NOM-INDEF
a *house*

This is the description of standard Arabic, but if we look at what happens in the spoken varieties, the situation changes considerably. Looking at the spoken varieties, considered in this study, we find that all of them use, to different extents, the indefinite article '*wahid'*, which is the numeral form for 'one'

[33] Jaber (2014) questions the traditional description of the use of definite in Arabic and the impossibility of a generic interpretation of Arabic bare nominals.
[34] Refer to the discussion presented in Lyons (1999:91-94).

(inflected in masculine and feminine). Moroccan and Syrian also employ the specific indefinite article '*ši-'* ('some') (Brustad, 2000; Harrel, 2004; Maas & Prochazka, 2022). Tunisian also uses the numeral for 'one' with the function of the indefinite article '*fard*' (Marcais, 1977; Mion, 2009; Turner, 2021). The spoken varieties lack the nunation and have therefore lost the case distinction as well. The nominal thus remains bare and, in that case, gives rise to an indefinite reading (Fassi Fehri, 2005). While in MSA bare nominals and indefinites are indistinguishable (Fassi Fehri, 2005; Salem, 2010), in spoken varieties, the situation is not so straightforward. Spoken Arabic comes to have a bipartite system of encoding indefiniteness: bare nominal or overt marking of indefiniteness.

It should be pointed out that in Egyptian and Syrian the use of '*wahid'* is restricted to human nouns, while in Moroccan the use is extended to non-human and even non-animate referents. In Moroccan, moreover, the numeral for one is always followed by the definite article '*wahid l-* '.

We follow Bustrad (2000) in asserting that the indefinite article occurs only with specific indefinites and is a marker of new topic, since it is used with non-identifiable referents, in first mention, that have relevance in the following discourse. For nonspecific indefinites, the bare nominal is used. Consider the following example, in Syrian Arabic, from Bustrad (2000:25):

(57) *fi* ***wahid badwi****, fat 'a l-mat'am. qal-lu la-l-garson, intini* ***buža***

there-is ***one bedouin****, entered-he into the-restaurant. said-he to-him to-the-waiter, give-me* ***ice-cream****.*

there's ***a bedouin*** *who went into the restaurant. He said to the waiter, give me* ***ice cream.***

The bedouin is the new character introduced for the first time and therefore with overt coding, while the ice cream is nonspecific and therefore realized as a bare nominal. The bare nominal also occurs in predicative functions and in idiomatic phrases (Harrell, 2004).

Moroccan and Syrian Arabic also have a prenominal bound article '*ši-'* /, used with both singulars and plurals, which occurs with specific referents and carries a meaning of uncertainty,

vagueness or potentiality (Harrell, 2004) e.g., *ši-razel* (a man, a sort of man, some man or other), *ši-ktab* (a book, some sort of book, some book or other).

In the description of Egyptian Arabic, the novel article *kida* is often mentioned, which is actually an adverb that becomes a marker of specific indefiniteness, indicating a referent that is not identifiable to the listener (but is identifiable to the speaker) (Brustad, 2000: 30; Turner, 2021: 9,12). Egyptian is the variety that has a more reduced number of definiteness and specificity markers, compared to other dialects[35].

We can conclude that, for what concerns the focus of this research here, all the varieties have some explicit element that marks the identifiability and non-identifiability of the referent.

As far as the expression of possession is concerned, Arabic belongs typologically to DG languages (see 1.4.4). The possessive is a suffix postposed to the noun, which by itself induces a definite reading and is never needed with the article. Like other Semitic languages, Arabic also has a particular possessive structure, known as the construct state. When a noun is modified by a possessive, that noun is said to be in the construct state: the first noun is the possessed, the other is the possessor. The former is in the bare form, while the latter can be definite (with the article) or bare. The possession relation induces a definite reading and thus prevents the possessed from carrying the definite mark (Lyons, 1999: 92-93).

Spoken Arabic knows both the construct state and possessive suffixes, but makes wider use of exponents[36] to convey the possession relation[37]; observe the following example (Brustad, 2000:75):

(58) *mal'ab it-tinis* ***bita'*** *in-nadi.* (Egyptian)
court the-tennis ***GEN****. the-club.*
the tennis court ***of*** *the club.*

[35] See Bustrad (2000:30-31), Badawi and Hinds (1986:25), for a discussion of formal means of expressing specificity in Egyptian Arabic.
[36] Harning (1980:10) defines the exponents as a "dialectal innovation". See Bustrad (2000:72) for a complete list of all the exponents.
[37] Arabic does not possess a verb that can be equated with the English verb '*to have'*, so, the construction of the type "X has Y" is normally expressed with the use of so-called pseudo verbs '*li*' and '*min*' (formally prepositions). See Bustrad (2000) for a description.

As can be seen from the example, the exponent '*bita"* cooccurs with the definite article and can also do so with the indefinite article. This type of structure brings spoken Arabic closer to Italian, in which the possession marker and the (in) definiteness marker cooccur.

Regarding the occurrence of the definite article with demonstratives, see the example below (Lyons, 1999: 119):

(59) ***hada al-****bustanu* (MSA)

this the*-garden*

this garden

Example (59) shows that Arabic (standard and spoken) belongs to the group of languages in which demonstratives and articles can occur. The system of demonstratives in spoken Arabic is complex, and here it will suffice to note the presence of the so-called "demonstrative article." These are demonstratives that have no deictic value or distance opposition, are unstressed, unmarked by gender, and are prefixed to the noun. They perform only anaphoric function but are distinguished from a real article because they cannot be used with uniquely identifiable referents that have not already been introduced into the discourse (Croft, 1990). See example (60) from Brustad (2000:118):

(60) ***had n****-nas d zman* (Moroccan)

this the*-people GEN. old*

those people of long ago

As with the other demonstratives, there is a mark of definiteness between anaphoric prefix and noun.

4. Chinese

Mandarin Chinese[38] belongs to the languages that do not possess an overt marking of definiteness and therefore resorts to other formal strategies to encode the opposition between identifiable and non-identifiable. The readings to which an NP can be subjected in Chinese are: definite, indefinite, or indeterminate (when the NP is neutral to the interpretation of identifiability) (Chen, 2015).

The formal strategies that Chinese employs to encode pragmatic definiteness are mainly of three types: (Chen, 2004): lexical, morphological and positional.

4.1 Lexical devices

In addition to pronouns and proper names, Chinese uses three main groups of determiners to mark a noun as identifiable: demonstratives, possessives, and universal quantifiers.

The most important demonstratives in Chinese are the proximal '*zhe*' (this) for singular and '*zhexie'* (these) for plural, and the distal '*na'* (that) for singular and '*naxie'* (those) for plural. Singulars can also take the form '*zhei'* and '*nei'*, in contemporary Chinese, especially in the variety of the capital. They perform all the functions of demonstratives: situational deictic, discursive deictic, anaphoric (contrastive). When usage is anaphoric, distal is preferred if the referent is referred to much later in the discourse, after the introduction of other referents, while proximal is preferred when the referent has just been introduced. Gundel et alii (1993:284-285) classify distal '*nei'* as the linguistic form that in Mandarin Chinese correlates most strongly with the cognitive state of "uniquely identifiable", which in English and Spanish corresponds to definite article. "In Chinese by contrast, this status [uniquely identifiable] appears to be sufficient for appropriate use of the distal demonstrative determiner *nei*; but Chinese apparently has no determiner which requires the referent to be familiar, but not necessarily activated. Thus, with respect to cognitive status requirements, *nei*

[38] This section will be devoted to the description of the formal means used for expressing definiteness in Mandarin Chinese, standard variety (see Arcodia & Basciano, 2020 for a short sketch of other varieties of Chinese).

behaves more like the definite article in English and Spanish than like the distal demonstrative in these languages". See the example from Gundel (1993:285):

(61) *Zuótian wǎnshàng wò shuì-bù-zháo. Gébi-de* ***nèi*** *tiáo gòu jiáo de lìhai.*

yesterday evening I sleep-not-achieve. Next-door ***that*** *CL dog bark ADV extremely.*

I couldn't sleep last night. **The** (lit**. 'that'**) dog next door was barking.

According to native informants, interviewed by the authors, the demonstrative would be appropriate even if the listener did not know at all that the speaker's neighbor has a dog (see 1.2.2).

What Gundel et alii (1993) noted supports observations already known in the literature (Li & Thompson, 1981) that demonstratives in Chinese would be the best candidates to become definite articles, following a well-known diachronic trajectory attested in other languages of the world (Greenberg, 1978). However, Chinese demonstratives have only begun the process of grammaticalization and are far from reaching the stage of proper articles. Most of the examples in which the deictic force of demonstratives is weakened and their use is close to the definite articles, are found in anaphoric contexts (Chen, 2004). Jenks (2018) notes that the demonstrative is used as a preferred device in anaphoric contexts, except if the referent is subject.

The use of demonstratives where the source of identifiability is general knowledge or frame-activated association is very rare, and limited to the vernacular style (Chen, 2004).

Regarding possessives, Mandarin Chinese falls under the label of AG language (see 1.4.4). Indeed, it is possible for a Chinese possessive to be separated from the head noun by an indefinite marker, which is not the case in languages such as English and Spanish. However, Chinese possessives cannot be completely assimilated to those of clearly AG languages such as Italian. In Italian, as already mentioned, the interpretation of the NP with the possessive depends on the presence of the definite or indefinite article. In Chinese, on the other hand, interpretation depends only on the indefiniteness marker '*yi'*, which if present induces a clear reading of no identifiability. Identifiability

of the referent emerges only as a result of the absence of this marker (Chen, 2004: 1157), as shown in (61):

(61) *Zhe shi wo-de **yi ge pengyou** gaosu wo de.*

*This be **my one CL friend** tell I DE.*

***A friend of mine** told me this.*

Chinese does not have a mandatory indefinite article, but the numeral for one '*yi'* now has a distribution very similar to what '*un'* and '*a*' have in Italian and English (Gundel et al., 1993). Here again we are facing a path of grammaticalization, which seems even more advanced than what is happening with demonstratives (Chen, 2003, 2015; Li & Thompson, 1981;).

'*Yi*' is a numeral and like all numerals in Chinese is followed by a classifier, but it has begun to serve some of the main functions of an indefinite article and also other uses that have not been recorded for indefinite articles in other languages. The functions served by *yi* + CL + noun are: numeral (62a), a presentative marker (62b), a marker of non identifiable specific reference (62c), a marker of nonidentifiable non specific reference (62d), and generalized article[39] (62e). (Givón, 1981; Hein, 1997). See examples below (Chen, 2004:1159-1160).

(62) *a. Wo zhi yao **yi zhi pingguo** jiu gou le.*

*I only want **one CL apple** then enough CRS.*

*I only want **one apple.***

*b. **Yi zhi xiao qi'e** yaoyaobaibai zou le shanglai*

***one CL little penguin** swaying walk PFV up*

***A little penguin** was waddling up*

[39] Those presented by Chen (2004) as uses correspond to the stages of the grammaticalization process of the indefinite article as described by Heine (1997). "Generalized article" denotes the last stage of the process, in which the article appears before all types of nouns and is no longer restricted to singulars but is also used with plurals and mass nouns.

c. Zhe jian shi wo zuotian qing le ***(yi) ge ren*** *lai.*

this CL issue I yesterday invite PFV ***one CL person*** *come.*

For this issue I invited ***a person*** *here yesterday*

d. Gankuan qu zhao ***(yi) ge ren*** *lai, shenme ren dou xing.*

Hurriedly go find ***one CL person*** *come any person all fine*

Hurry up and get somebody; anybody will be just fine

e. Ta kan shangqu xiang ***(yi) ge faguoren***

he look up like ***one CL Frenchman***

He looks like ***a Frenchman***

Actually, in non-specific indefinite contexts, the bare nominal is more common than the construction with '*yi'*. This is in line with the trend whereby the more nonreferential an NP is, the more likely is it to drop the indefinite marker '*yi*' and the classifier (CL) and assume the form of bare NP (Chen, 2003, 2015). Chinese also possesses complex indefinites: pronouns and indefinite determiners that will not be devoted space to in this chapter. [40]

4.2 Morphological devices

One morphological strategy Chinese can resort to in order to express (non) identifiability of the referent is reduplication. Monosyllabic classifiers or monosyllabic nouns undergo this process (Chen, 2004). When in preverbal position these obtain distributive meaning (Paris, 2007). See example (63) from Chen (2004), and they cannot appear in indefiniteness inclined positions (see next section):

[40] On this aspect, refer to Chao (1968), Chen (2004: 1162-1163).

(63) ***Zhong-zhong*** *yinsu* *dou* *dei* *kaolu* *jinqu.*

CL *factors* *all* *must* *consider* *in*

Every *factor has to be taken into consideration*

4.3 Position

As already seen, Chinese does not have obligatory determiners, and in fact admits bare nominals even in argument position (Chierchia, 1998). When the formal devices described above are not applied, and thus the noun is left bare, the identifiability interpretation is normally indicated by the position of the NP in the sentence. Word order is certainly the most used formal strategy in Chinese to encode the identifiable/non-identifiable opposition (Givón, 1990).

Some positions show a strong inclination for a definite reading, while others tend to induce an indefinite reading. Importantly, these are not restrictions but tendencies: the inclination is manifested in terms of higher text frequencies of the nominal expressions of definite versus indefinite encodings occupying the particular position in sentences (Chen, 2004, 2015):

(64) **Definite-inclined positions:**
Subject
'*ba'* object[41]
preverbal object
first object of ditransitive sentence

(65) **Indefinite-inclined positions:**
object of the presentative verb '*you'*
postverbal NP in presentative sentences
postverbal NP in existential sentences
second object of ditransitive sentences

A NP coded as identifiable through a lexical or morphological strategy must occur in a tendentially definite position (Chen, 2015). In summary, it can be said that, in Mandarin Chinese, preverbal

[41] The construction with '*ba'* in Chinese has a clear structure: S + ba + O + V. What the communicative function of this construction is and what kinds of verbs may occur in it is a complex and much discussed question. For the purposes of this chapter, it should be stressed that a *ba*-NP refers to something that the speaker believes hearer knows (Li & Thompson, 1981: 463-491).

subjects (66a) and objects (66c) tend to be definite. See examples below (Li & Thompson, 1981:20-21). Given the general tendency in many languages for topics to be definite (as given or shared information), it should be emphasized that in Chinese, topics and subjects align, and their unmarked position is preverbal (given Chinese's canonical SVO order), thereby resulting in the preverbal position tending towards a definite reading. This applies to objects as well when they are topical (yielding a marked SOV order).

(66) a. ***rén*** *lái le*	b. *lái le* ***rén*** *le*	c. ***shu*** *wo mai le*
person *come PFV*	*come PFV* ***person*** *PERF*	*book I buy PERF*
the persons *have come*	***some persons*** *have come*	***the book****, I bought it*

In other words, in Chinese the first element of a sentence is not necessarily a subject, rather the topic of the sentence, that is, what is being talked about (Li & Thompson, 1981). From the cross-linguistic literature (Givón, 1984, 1990) the strong correlation between the syntactic role/position of subject, the notion of topic and definite reading clearly emerges. What is being talked about is generally the given element, i.e., known to the hearer (and therefore identifiable) which is usually also the subject of the sentence. For this type of element, the preverbal position is reserved in Chinese. It follows that in Mandarin bare nominals in preverbal position cannot be interpreted as indefinites (Cheng & Sysbema, 1999) (66b), and that indefinites are therefore bad subjects (Sparvoli, 2017).

This limitation of indefiniteness in preverbal subjects is shared by Chinese and Italian as far as bare nominals are concerned. Indeed, it has been noted that in both languages zero-determiner nouns, when subjects, can obtain indefinite reading only when in postverbal position (Longobardi, 1994; Sparvoli, 2017).

If we look at (66b) we notice that in Mandarin it is still possible to have indefinite subjects, which are normally placed postverbally. The phenomenon of subject-verb inversion, with respect to the canonical order (VS instead of SV), responds to the need to put given information before the new

information. In these cases, the new element is the subject. This type of order occurs in '*you'* (exist/to be there/to have) constructions (67a), other kind of existential sentences (67c) and presentative sentences[42] with inaccusative verbs (67b and 66b) (Lyons, 1999). This type of structure in which the inversion of the canonical order is required has been observed in the typological literature in other languages (Beaver et. alii, 2006) and has a parallel in Italian as well (Sparvoli, 2017). Note the following examples from Li and Thompson (1981:509-514):

(67) a. *Chéng li <u>yǒu</u>* ***gōngyuán***.
city in exist ***park***
there are ***parks*** *in the city*

b. *Lái le* ***yí ge kèrén***
come PRF ***one CL guest***
a guest *is coming*

c. *qiánmian shì* ***yi ge huayuán***
in:front be ***one CL garden***
what's in front is ***a garden***

As for the object, its canonical position is postverbal, and it therefore tends to receive indefinite reading. See the example from Li and Thompson (1981:96):

(68) z*hèi jian shì nib u néng guang máfan* ***yi ge rén***
this CL matter you not can only bother ***one CL person***
this matter, you can't deal with it by bothering only ***one person***

As with subjects, it is possible to move the object before the verb, assign it the position reserved for the topic and interpret it as definite. Indeed, it has been seen that both the object preceded by particle '*ba',* and preverbal objects show traits of high topicality (Chen, 2004). In these cases, the object is

[42] We follow Li (2014) in differentiating existential and presentative sentences. "Existential constructions designate stative situations; they are topic-comment in nature. In narrative discourse, they actively participate in various types of background descriptions. Presentative constructions introduce new entities into discourse; they designate bounded dynamic events". The first type is related to the presence of a locative phrase, the second of a verb of motion (Sparvoli, 2017).

presented as identifiable. Consider the example (66c) for the preverbal object and the following one for the '*ba'*object (from Li & Thompson, 1981:464)[43].

(69) *ta* ***ba fanting*** *shoushi ganjing le.*

s/he ***BA dining:room*** *tidy clean PFV.*

s/he tidied up ***the dining room****.*

All this also applies to NPs with numerals (numeral + CL + noun in preverbal position is definite subject, if it follows *'you'* is indefinite subject, etc.) (Sparvoli, 2017). Even though a corpus-based study by Chen (2004) showed that NPs with cardinals and those with the quantifier '*ji'* (several) are often neutral to the definite/indefinite contrast.

4.4 Indeterminate reading

In Chinese, there are bare referential NPs whose interpretation with respect to the notion of identifiability is neutral. Consider the following example from Chen (2015:411):

(70) *Zuowan lianyouchang da huo,* ***tie men*** *dou shao hua le.*

last:night refinery big fire ***iron gate*** *even burn melt CRS.*

there was a big fire at the refinery last night. ***The/an iron gate/gates*** *melted in the fire*

The bare noun '*tie men*' (iron gate) in (70) can refer to an identifiable referent, if there is only one iron gate at the refinery, which is expected to be known to the addressees as part of their common background knowledge or as a frame triggered entity. The bare NP is also appropriate when the

[43] Because these are trends, exceptions to these definiteness effects are also found. See Li & Thompson (1981:167-168) for the exceptions to indefinite subjects in preverbal position and Chen (2004:1170) for an explanation of cases such as (62b). See also Chen (2004:1169) for a discussion of definite subject in existential or presentative sentences.

speaker cannot assume the addressees are aware of how many iron gates there were to the refinery and which one or ones were melted in the fire. And it is likely that the speaker is similarly unaware of this. The only information that the speaker wants to convey to the addressees is that the fire was so damaging that it melted one or more iron gates. In the latter scenario, the statement refers to an unidentifiable object. Also going under the indeterminate category are those nominals that in languages such as Italian and English would be formally definite, while indicating referents with partial identifiability (weak definites in other terms) (Chen, 2004). The interpretation of indeterminate expressions with respect to identifiability is expected by the speaker to be inferred by the addressee, based on topicality, position, the availability of an identifiable referent in the context that meets the descriptive content of the expression, as well as other relevant information of the utterance in the universe of discourse (Chen, 2004).

Following Chen (2015), it can be concluded that definiteness as a grammatical category is not fully developed in Chinese, so speakers resort to other formal strategies to encode semantic-pragmatic definiteness.

5. Summary and predictions

In this section, a table is added (Table 1), in which the main features of the four linguistic systems are summarized and compared, and hypotheses pertaining to cross-linguistic effects relevant to the study presented in the subsequent chapters are derived.

In the following chapter, the predictions will be related both to a more precise analysis of cross-linguistic influence dynamics (see 3.3) and to the findings of research on bilingual acquisition (see 3.5).

Identifiable status	Linguistic Form	ITA	SPA	ARA	CHI
NON ID.	Bare NP	+ *(constrained*)*	+ *(constrained*)*	-	+ (*postverbal)*
	Numeral for 1 + NP	+	+	+	+
	Postverbal NP in presentative structure	+	+	+	+
ID.	Bare NP	+ *(constrained**)*	+ *(constrained**)*	-	+ *(preverbal)*
	Def. art. + NP	+	+	+	-
	Demonstrative + NP	+	+	+	+
	Def. art. + Demonstrative	-	+	+	N.A.

Table 1 Summary of the main definiteness-related features of the four linguistic systems considered in the study.

* The occurrence of non identifiable bare nouns in Italian and Spanish is semantically and syntactically constrained (see 2.1 and 2.2)
** The occurrence of identifiable bare nouns in Italian and Spanish is semantically constrained (see. 2.1 and 2.2)
N.A.: not applicable.

As emerges from the Table 1 above, the only area of divergence where the effects of Spanish on Italian can be expected is the co-occurrence of the article and the demonstrative in the same NP (although the pragmatically marked nature of the phenomenon in Spanish should be aknowledged).

As for Arabic varieties, areas where the effects of the family language might be evident include the co-occurrence of the article and the demonstrative in the same NP in anaphoric contexts. It is also expected that there would be less extensive use of bare nouns by Italian-Arabic bilingual children compared to monolingual peers, in contexts where Arabic prefers the use of the definite article while Italian allows bare NPs: mass nouns in PPs and bare plurals in non-identifiable contexts.

An effect of Chinese on Italian is expected in two forms; where Italian would use the definite article, Chinese would tend to :

- overuse demonstratives in an anaphoric contexts,
- omitt the definite article in identifiable contexts,

As for the canonical word order, it is the same both in Italian and Chinese: Subject-Verb-Object (SVO), with the subject typically being topical and definite. There is also substantial convergence between the two languages in presentative sentences, which feature a post-verbal subject in both languages (with '*c'è + NP*' in Italian and '*you + Cl + NP*' in Chinese, meaning "there is"' in English; see 2.4.3). Regarding the object, in Italian and Chinese, it is normally post-verbal. However, in Italian,

its position remains unchanged whether accompanied by the definite or indefinite article; whereas in Chinese, to receive a definite reading, it must move before the verb; otherwise, it will tend to raise an indefinite reading (see 2.4.3 above). From this, one might expect that cross-linguistic influence manifests at the level of the object, and that a Chinese-Italian bilingual speaker, when speaking Italian, would tend to place the object in pre-verbal position when they want to signal its identifiability to the listener.

3. Definiteness in bilingual acquisition

After illustrating the main encoding strategies of definiteness in the languages involved in the present study, the following chapter is dedicated to the acquisition of definiteness and its formal means by monolingual and bilingual children. It is structured as follows: in section 1, the development of determiners (especially articles) and discourse-pragmatic skills in monolingual children will be discussed. Section 2 will address the same topics in relation to bilingual children, while section 3 continues with the discussion of cross-linguistic influence and presents the debate on the subject. Section 4 is about non-linguistic factors that influence bilingual development. The chapter concludes with reporting the research questions and predictions, based on previous literature.

1. Acquisition of definiteness in monolingual children

A large number of studies have investigated the acquisition of nominal determiners, in particular articles, in monolingual children, in different languages and within different theoretical frameworks. One thing on which all studies converge is that articles emerge very early in children's language. Brown's (1973) pioneering study found that Adam, Eve and Sarah acquired the English article system between the age of 2;8 and 3;5. After Brown's work, several studies on English acquisition followed, taking different positions: in line with functionalist theories (Karmiloff-Smith, 1979; Pine & Lieven, 1997; Pine et al., 2013; Rozendaal & Baker, 2008 to name a few) or with Universal Grammar approaches (Radford, 1990; Valian, 1986, 1991; Valian et al., 2009; Wexler 2011 among others). Within these two approaches, comparative studies have also been conducted (see Bassano, 2015, for an overview). The early emergence of the determiner category in children is confirmed for various languages and with different rates, always within the third or fourth year. In particular, for Romance languages like Italian, the early emergence of the determiner category has been assessed (Bassano et

al., 2008, 2011; Gavarró et al., 2006; Guasti et al. 2008; Lleó, 2001 among others). In these languages, the acquisition of determiners is a central process for grammar development and the mastery of various discourse functions. In spoken productions, one of the first steps that children achieve, relatively early, before the age of 3, is precisely the use of determiners with nouns in obligatory contexts, instead of producing bare nouns. Despite this early acquisition, Karmiloff-Smith (1979) showed that the development of the determiner system is gradual, due to its plurifunctionality, as children acquire the various functions of determiners progressively and endow these markers with plurifunctional status only late. Mastery of correct and appropriate use of determiners can take a long time to achieve, with the acquisition process lasting into the primary school years or later.

For Italian, Chierchia et alii (1999) argue that in the acquisition of determiners, children go through three stages:

1. BARE-NOUN STAGE[44], where all nouns are bare.
2. VARIATION STAGE, where there is a mix of nouns with and without determiners.
3. TARGET STAGE, where determiners are present when required.

Italian children seem to acquire the determiner category very early, sometimes even before 1;6 years and never after 2 years (Pizzuto & Caselli, 1992). In Italian, children generally cease to omit articles in obligatory contexts before they turn 3;0 (Caselli et al., 1993; Pizzuto & Caselli, 1992). According to Chierchia et al. (1999), children acquiring a Romance language stop omitting articles at a lower mean length of utterance (MLU), compared to children acquiring a Germanic language (see also Guasti et al., 2008). However, Kupisch (2007) underlined that interpretations based solely on MLU should be approached with caution. Given that nouns make up a substantial part of early childhood lexicon, the increase in MLU is not entirely independent of article usage. The author rephrased

[44] See Pérez-Leroux & Battersby (2009) for a discussion on the bare-noun stage (in Spanish child-language).

Chierchia's formulation and argued that "children who use articles more frequently than others also have higher MLUs" (2007: 62). The factors involved in the process of determiner acquisition are diverse, and various studies have emphasized the importance of one or the other: prosodic, lexical-semantic, morpho-syntactic or pragmatic factors (Bassano, 2015). It is not possible to delve into the details of single proposals here; only a few will be mentioned as examples.

Various studies have considered phonological and prosodic factors in accounting for the realization of grammatical morphemes like determiners in children (see Bassano 2015: 28-33; Demuth, 2011 for reviews). Different accounts have been proposed, such as the Trocaic Template Model (Gerken, 1994, 1996) or the Prosodic Licensing Hypothesis (Demuth, 2011; Lleó, 1997; Lleó & Demuth, 1999), to name a few. The first model suggests that children's early words tend to take the form of trochaic bisyllabic feet Sw (Strong-weak), making them more likely to include grammatical morphemes like articles if they can be prosodified as part of a foot. The second model expands on the first and aims to account for cross-linguistic variations, linking them to how grammatical morphemes are prosodified in the target language and the extent to which input provides children with prosodic patterns at the lexical level. A highly frequent syllabic structure in the target language can facilitate the acquisition of a determiner, when the determiner and a noun result in an NP with that highly frequent syllabic structure, which the child is already familiar with (for Italian, see Crisma & Tommasutti, 2000; Giusti & Gozzi, 2006).

From a syntactic perspective, some authors working within the generative frame, argue that, initially, a child's grammar includes only lexical categories, with functional categories like determiners appearing later and being constructed through a bottom-up process (Hulk, 2004; Radford, 1990). This may explain the presence of many bare nouns in the early acquisition phase. Another generative proposal suggests that the child's grammar has both lexical and functional categories from the beginning, and that article omissions result from processing limitations (Avrutin & De Lange, 2004; Baauw et al., 2002; De Lange et al., 2009). Usage-based scholars, on the other hand, tend to believe that determiners are acquired in an item-by-item process, as children do not generalize the

use of different determiners on lexical nouns but use fixed determiner-noun combinations (Lin & Li, 2022; Tomasello, 2003). Other scholars have emphasized semantic-pragmatic explanations, such as sensitivity to specific/non-specific distinctions (Roeper, 2006). According to this hypothesis, children omit determiners primarily for non-specific referents and gradually progress to full NPs with both definite and indefinite determiners as their semantic sensitivity develops. Other studies argue that children omit more articles in contexts where the referent is highly salient, initially relying more on the non-linguistic, perceptual context to make their reference clear (Baauw et al., 2002). Rozendaal (2008), studying referential expression acquisition, including determiners, in Dutch, English, and French children, underscored the strong interplay between morphosyntax and pragmatics. Children would start to associate determiners with pragmatic functions from the moment that the form have become productive, irrespective of age. Acquisition of a morphosyntactic form might trigger its correct use for pragmatic function (2008:267). The author also argues that this proposal applies mostly to the factors of specificity and new/given in discourse, for which the cognitive basis is assumed to be already present before the start of determiner use (2008: 273).

This brief list of explanatory proposals reveals that the acquisition of determiners is a composite process determined by the intervention of various factors.

1.1 The acquisition of definite and indefinite distinction

In this subsection, I will specifically focus on the acquisition of the crucial distinction between definite and indefinite articles, which is central to the present work.

Despite significant differences in participants' age, data collection methodology, and theoretical-interpretative frameworks, most studies have highlighted that children face challenges in adequately using definite and indefinite determiners. By "facing challenges" it is meant that children's productions include errors related to the appropriateness of the selected determiner in the given

context. The types of errors found in the literature regarding the use of articles fall into three categories (without considering errors in agreement between the noun and the determiner):

- omission of the determiner in obligatory contexts
- use of indefinites in definite contexts
- use of definites in indefinite contexts

As shown above, the transition beyond the omission phase occurs very early, especially in children who speak Romance languages. In the literature, the use of indefinites instead of definites has been categorized as “incoherence errors” (Emslie & Stevenson, 1981) or “discourse integration errors” (Krämer, 2003). Substitution errors involving the use of definite articles in place of indefinite ones have been described as “egocentric” (Maratsos, 1976), indicating a struggle to consider the hearer's perspective when introducing a new referent known only to the speaker.

One specific tradition links children’s errors in definiteness to an egocentric stage in their cognitive development (Karmiloff-Smith, 1979), or its expression as an immature pragmatic principle (Schaeffer & Matthewson, 2005). This suggests an expected correlation between the occurrence of definiteness errors and the relevant Theory of Mind[45] component (Gundel, 2009; Schafer & de Villiers, 2000).

In his early study, Maratsos (1976) explored the usage of definite articles in 3-to-4-year-old English-speaking children within two distinct contexts: anaphoric context and bridging context. The author found that three-year-old children correctly employed definite articles in the anaphoric condition at a rate of 55%. By the age of four, children had mastered anaphoric article use, exhibiting an approximately 95% production rate. For the bridging context, three-year-old children correctly

[45] Theory of Mind (ToM) involves the cognitive capacity to attribute mental states (like beliefs, desires, etc.) to others and it is established around age 4. Second-order ToM, which entails reasoning about other people’s mental states regarding additional mental states, develops at a later stage, typically around the age of 6 (De Cat, 2015: 273-274; Tomasello, 2018). See Camaioni, 2022 for an overview on ToM development.

used the definite article 83% of the time, and four-year-old children demonstrated a 98% accuracy rate.

Schafer and de Villiers (2000) later confirmed the dissociation between anaphoric and bridging use of definite articles in a study involving English-speaking 3-to-5-year-old children. Notably, there was considerable between-group variability unrelated to age, as some five-year-old children performed worse than their three-year-old counterparts. As for type of errors, in both conditions, children primarily made omission errors, accounting for approximately 8-13% in the bridging condition and 23-40% in the anaphoric condition. Additionally, children occasionally substituted the definite article with the indefinite one, more frequently in the anaphoric condition compared to the bridging condition.

Regarding bridging contexts Avrutin and Coopmans (2000) argued that children can compute these bridging relationships (defined as part-whole relations) based on world knowledge and are able to establish this kind of reference early in various languages. In their work, they noticed that younger children show a chance distribution in the use of articles in the bridging context and suggested that this might be due to insufficient processing resources required for these computations, which involve both syntactic and extra-syntactic knowledge (Avrutin, 1999).

Schafer and de Villiers (2000) have proposed that problems with Theory of Mind (TOM) and with the [± hearer] distinction are the main causes of errors in children's productions. The difficulties encountered in using definite articles anaphorically have been ascribed to children's difficulty in assessing other people's belief or perspectives and grasping the [± hearer] distinction. Conversely, in the bridging condition, children demonstrate proficiency in establishing uniqueness. This theoretical explanation predicts that the definite articles used by young children lack the target adult-like semantic features, because early definite articles lack the [+hearer] feature associated with Theory of Mind (TOM).

Gundel (Gundel et al., 2007; Gundel 2009; Gundel & Johnson, 2013) argues that errors with articles in children's productions stem from their lack of sensitivity to Grice's Maxim of Quantity,

which is connected to Theory of Mind. According to Gundel and Johnson (2013), children, at least up to the age of 4, may not be sensitive to the quantity of information relevant in the context of an utterance. The authors report the experiment by Schaeffer and Matthewson (2005), where a mouse drawing a house was showed to the participants. When asked "What was the mouse drawing?", all adults answered "a house", while half of the children responded "the house". Gundel and Johnson comment on the experiment, asserting that children provide an appropriate response if the shared knowledge between them and the experimenter (the image of the mouse was visible to everyone) is considered, but that the answer was not adult-like. Children's use of the definite article would be inconsistent with the pragmatic fact that this form provides more information than is necessary in the context of the experiment. For adults, the indefinite article is used when the conditions for using the definite article are met, but unique identifiability is irrelevant, as in the case of the experiment (Gundel & Johnson, 2013:55). Children behavior would be linked to not fully developed Theory of Mind (ToM), which does not enable children to represent the whole set of epistemic mental states of the interlocutor in order to understand the amount of information required/necessary in the context (Gundel, 2009; see also Rubio-Fernandez, 2021).

However, there is a large corpus of research suggesting that children are indeed sensitive to pragmatic principles of reference from a very early age (*contra* Modyanova & Wexler 2007; Schaeffer & Matthewson, 2005; Valian, 2013). Skarabela and Allen (2013)'s study examines a naturalistic corpus of spontaneous Inuktitut speech by four children (aged 2;0–3;6) to explore how they realize new referents. Contrary to arguing for pragmatic deficiencies, the study suggests that children strategically tailor their messages for the interlocutor, adhering to the Gricean Maxim of Quantity (see also Bredart, 1987). Young children also display sensitivity to information structure, using overt forms more frequently for new referents compared to old ones (De Cat, 2011). Nevertheless, preschoolers produce a number of definiteness errors, also after the essential linguistic distinctions are established (such as definiteness, topicality, and so on). Other types of factors, than pragmatic deficiencies, must be invoked to explain the presence of errors with articles.

De Cat (2011, 2013) leans towards other kinds of cognitive limitations. She contested that the difficulties with articles in children are merely linguistic, as evidenced by the results of the bridging conditions, where no or very few substitution errors occurred. This alignment with adult-like grammar supports the view that challenges with anaphoric article use stem from discourse integration difficulties, possibly due to the high processing resource demands while tracking discourse reference (Serratrice, 2006). This would also explain the substitution of the definite with an indefinite, which does not necessarily reflect linguistic issues, but rather cognitive challenges in evaluating the newness status of referents. This view contrasts with Schafer and de Villiers's (2000) perspective, where children's grammar is considered incomplete due to a lack of the [+hearer] feature, in fact, children linguistic knowledge would be not incomplete, but adult-like (De Cat, 2011:858).

De Cat (2013)'s study was the first to specifically address the correlation between Theory of Mind (TOM) and egocentric errors in preschool children, demonstrating that no correlations emerged. Children's performance in her experiment, which aligns with findings from other studies (De Cat, 2009; Gundel, 2011), suggests that the linguistic foundation of discourse competence, specifically the understanding of rules governing definiteness choices based on referent givenness, is established before the full development of Theory of Mind. This does not imply that TOM does not play a role in children's referential abilities; rather, its influence is indirect. It should be noted that TOM is a complex and multidimensional construct, and executive functions appear crucial for its emergence and development. De Cat (2015) focuses precisely on executive functions[46] as a possible explanation for children's over-use of forms and as the cognitive underpinning of referential abilities in children. Particularly in the production of lengthy texts, working memory appears to play a fundamental role in referentiality. Whitely & Colozzo (2013) investigated the relationship between working memory and the adequacy of referential expressions in children (5;5-8;7) during a narrative task, categorizing each expression according to their referential function: introduction or maintenance. The findings

[46] Executive functions have been categorized into three distinct components: suppressing undesirable responses (inhibition), shifting between tasks and mental sets (flexibility), and updating and monitoring working memory representations (De Cat, 2015:264; see also Friedman et al., 2000; Miyake et al., 2000).

revealed a moderate positive correlation with the maintenance function, but not with the introduction function. The authors argue that introduction proved challenging for many children, revealing their difficulties to manage the discourse model effectively. The results suggest that, for the demanding introduction function, basic memory capacity played a vital role, while updating had a more substantial influence on maintenance, particularly for younger children.

De Cat (2015) argues that inhibitory function too plays a significant role in referential choices, since it is necessary to inhibit one's own perspective and to adopt that of others.

If this line of interpretation is correct, referential disruptions are found to result from cognitive limitations, including young children's challenges in maintaining continuity between events or pictures (related to general story-telling abilities). Additionally, their tendency to presume a broader common ground with their conversational partner, potentially intensified by the assumption of joint attention, contributes to these disruptions. Children appear more inclined than adults to assume a shared perspective with their addressee and are less vigilant than adults in monitoring the need for perspective adjustments. It has been argued that discourse integration and egocentric errors may obscure adult-like linguistic knowledge that underlies definiteness choices for encoding information status in children (De Cat, 2013). Children do not seem to have difficulties in creating an abstract discourse model; rather, they exhibit challenges in handling it: maintaining, updating, correcting, and adapting it to the interlocutor. These are complex and demanding cognitive operations, even for adults, as they require flexibility and maturity in executive functions. It is at this level that differences between children and adults would become apparent and have linguistic implications.

1.2 Errors with articles as speech-production errors

In the explanatory framework just described (De Cat, 2011, 2013) errors with articles have been considered as performance-based errors that do not accurately reflect the children’s linguistic abilities. They are attributed to children’s challenges in linguistically integrating a previously mentioned entity,

often influenced by experimental settings (De Cat, 2011) or the demand of processing (Serratrice, 2006). This is in line with speech-production models that link referential failures to limitations in the process of language production itself (see Adams & Gathercole, 2000; Levelt, 1989). Several studies share the assumption that linguistic and cognitive abilities (working memory, executive functions) mobilized during production are, to some extent, linked to referential adequacy and can explain the increase of adequate uses over development (see Grigoroglou & Papafragou, 2019; Matthews et al., 2018, for recent reviews).

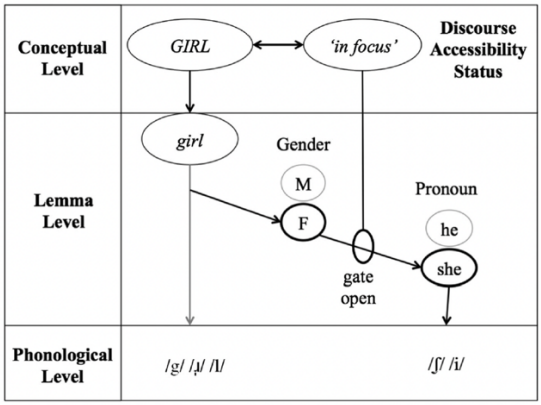


Figure 1 Arnold and Zerkle (2019)'s adaptation of the model of reference production from Schmitt et al. (1999).

Speech-production models have been developed in adult's psycholinguistics, starting with the work of Levelt (1989). The Levelt's model (1989) explains spontaneous speech productions and consists of three components: conceptualizer, formulator (with grammatical and phonological encoders as its subcomponents), articulator, speech-comprehension system, and monitor. For the focus of this chapter, I will concentrate on the first two components. The conceptualizer (or conceptual level) generates the message to be produced, converting the speaker's intentions into the so-called pre-verbal message. At this level, the evaluation of shared knowledge between the speaker and interlocutors occurs, and, in the case of nominal reference, the speaker assesses the identifiability of the referent by listeners. This is where the discourse model is formed. Each referent is then assigned an accessibility index, informing the interlocutor if the referent is identifiable and where it can be

found: in encyclopedic knowledge, perceptual context, the preceding discourse, etc. The accessibility index is thus part of the pre-verbal message of the noun; there is no separate encoding for the semantic information of an article at this level (De Lange et al., 2009). The selection and production of the article will be triggered in successive steps by the identifiability status assigned to the referent in the pre-verbal message of the noun, according to language specific rules.

The output of the conceptualizer serves as the input for the formulator, which transforms the pre-verbal message into a linguistic structure through two consecutive steps: grammatical and phonological encoding. In the grammatical encoder, the semantic information of the pre-verbal message triggers the activation of the corresponding lemmas in the mental lexicon (lemma level). The lemma encompasses the meaning, syntactic properties, and grammatical features of the lexical item. This results in a lexical-based model where the lemmas are "the driving force behind the speaker's construction of the surface structure. It is in the lemmas of the mental lexicon that conceptual is linked to grammatical function" (Levelt, 1989: 162). The syntactic category of the lemma then initiates category-specific functional procedures to build syntactic structure: NP in our case. The grammatical encoder utilizes the information from the lemma to choose a suitable article (or other specifiers; Levelt, 1980:170) that conveys the identifiability status (as indicated by the accessibility index) of the intended referent and creates the surface structure of the NP. The surface structure forms the foundation for accessing the phonological forms of words and constructing a phonetic as well as an articulatory plan for the utterance to be produced.

Arnold and Zerkle (2019) present the referential production model of Schmitt et al. (1999), adapting it to exemplify the production of a pronominal form (*she*) in English. As depicted in Figure 1, operations such as the evaluation of common ground, of the information status of the referent, and other discourse-pragmatic assessments take place at the conceptual level, the level responsible for creating the message itself. Only later, at the lemma level, do these decisions lead to the selection of a linguistic form (a pronoun in the example) with its syntactic properties.

It would be at the conceptual level, therefore, that pragmatic errors in the use of articles would have their origin, potentially due to excessive cognitive load or immaturity.

What psycholinguistic studies on adults have highlighted is that operations such as assessing common ground, maintaining the abstract discourse model, and inhibiting one's own perspective are cognitively demanding even for adults and can lead to errors. To date, much evidence suggests that adult speakers do not perfectly track all sources of information about addressee knowledge and that they fail in referential production many times (Fukumura & Van Gompel, 2012)[47]. It is possible that children and adults use the same kind of information and encode it similarly, but with varying reliance on cues (from previous discourse or perceptual context) and different cognitive limitations (De Cat, 2013). As cognitive control improves, particularly towards the conclusion of the preschool years, children are likely to do better in tasks where the monitoring of their interlocutor's perspective is explicitly and consistently emphasized.

De Lange et alii (2009) also refer to Levelt's (1989) speech production model to explain the omission of articles in children, asserting that it is due to cognitive limitations.

The message of the conceptualizer is transformed into linguistic material through a morphosyntactic channel in the formulator, which is the most efficient route in healthy adult speech. Avrutin (2006) suggests that an alternative encoding route, based on presupposition left to the interlocutor, competes with the morphosyntactic route. This alternative route may take precedence if the morphosyntactic channel is not fully developed or impaired, as observed in child and aphasic speech. The authors (De Lange et al., 2009) propose that children's maximal syntactic channel capacity is lower than that of healthy adults. Consequently, elements requiring significant cognitive effort and processing time, like articles, may be omitted. In children, this limitation likely stems from the ongoing maturation of relevant brain structures and the resources needed for the automatic, rapid activation of lexical items. However, children do not omit articles all the time and, in fact, the

[47] See also Pickering and Garrod (2004) for a conversational speech model and repair strategies in common ground assessment in adults.

availability of processing resources in a given situation can be affected by other factors: e.g., tiredness, attention or resource-consuming nonlinguistic activities (De Lange et al., 2009).

1.3 Types of task and children's performance in article production

As seen above, cognitively demanding activities can lead to different performances in the use of articles by children; the type of task and the required cognitive resources are said to be decisive factors in the use of articles (De Cat, 2013; De Lange et al.2009).

Many studies have investigated the appropriate use of determiners, specifically articles, across the life span, often yielding conflicting results. One line of findings suggests that children can use articles appropriately in discourse from a very early age (De Cat, 2011), while another line of research argues that mastery of article functions continues into adolescence (Hickmann 2003a). It is not necessarily the case that these two perspectives are incompatible. The results imply that variability within specific age groups is greatly influenced by the task's nature (De Cat 2013). Generally, preschoolers exhibit better performance in familiar situations than in experimental settings (Ninio & Snow 1999) and with familiar items (e.g., sets of animals) rather than unfamiliar items (e.g., abstract shapes) (Yule, 2013). When engaged in a relatively simple and interesting task, children's language doesn't seem entirely egocentric. However, in certain situations, especially when faced with great cognitive, linguistic, and social demands, they may display egocentric behavior more frequently than older children and adults (Bryant, 2015).

Research asserting early competence in determiner use often examines spontaneous productions in familiar, uncontrolled situations. These studies focus on short utterances, particularly in question-answer sequences. In contrast, research emphasizing later mastery widens its scope to connected discourse, utilizing narrative production tasks, and highlights differences between adults and children enduring till later stages of development (Berman & Slobin 1994; Hickmann 2003b; Schmicke et al., 2015; Schmicke et al., 2020).

The nature of mutual knowledge in tasks strongly impacts children's use of referring expressions (Hickmann et al., 1995; Kail & Hickmann 1992). Dialogic productions typically involve mutually known entities (De Cat, 2011; 2013), while experimental studies (Chondrogianni & Marinis, 2015; Schafer & de Villiers, 2000) may focus on specific functions that differ from those necessary in contexts requiring integration, like cohesive narratives. Considering these differences, the protracted path of determiner use acquisition in narratives is explicable and aligns with the cognitively demanding nature of tracing referent identity. This process necessitates late-developing abilities such as memory retention, updating the discourse model, and effective linguistic expression, presenting a formidable task for children, even at school age (Berman, 2015).

Another aspect of the task to consider is the presence of visual stimuli. According to De Cat (2013) performances in using articles adequately are poorer in visual stimulus experiments due to reliance on visual over discourse context, assuming joint attention. Over-reliance on the visual context in picture-based experimental tasks may stem from difficulties maintaining an abstract discourse model, making the visual context a more reliable source for common ground evaluation. This allows for lightening the cognitive load, given that the ability to distance oneself from the immediate perceptual context is demanding in terms of executive function abilities (De Cat, 2013, 2022)[48].

1.4 Articles as elements of discourse cohesion

A factor contributing to differences in performance in the use of articles between tasks of spontaneous conversation and narrative text productions is that the latter requires children to use articles not only to signal the informational status of the referent but also as elements useful for creating textual cohesion. In other words, articles exhibit multifunctionality.

[48] See Keysar et alii (2000) for a similar interpretation in adults, where an egocentric perspective is said to function as a cognitive load reduction strategy.

When considering the role of cognitive complexity in the acquisition timing of various linguistic devices, including determiners, the key factor is children's ability in handling multifunctionality (Karmiloff-Smith, 1985). Levelt's (1989) characterization of reference production underlines this, highlighting that definite descriptions, among other grammatical devices (e.g., pronouns), serve as an anaphoric tool linking the interpretation of the current utterance to what was previously mentioned or what is yet to be said. Definiteness, contributing to discourse cohesion, would be akin to connector words like "because", "and" or "so" (Grosz et al., 1995). The function of textual cohesion creation is probably the most difficult for children to acquire. This idea is consistent with the results from various studies and would explain both the differences between definites in bridging and anaphoric contexts and the prolonged difficulties in narrative tasks (Arnold & Zerkle, 2019).

Late developmental progressions in narratives suggest that discourse-internal anaphoric principles are acquired after 6 years or even later (Bamberg 1986, 1987; Schmicke at al., 2015). Third person fictional narratives are especially challenging for children, since they require tracking the reference of one or more characters during the unfolding-discourse, moving from one referent to another and reintroducing characters through anaphoric connections, creating a cohesive and navigable text for the listener.

This kind of text also requires anchoring the actions of the characters to spatial elements and temporal references. Non-animated entities that serve as spatial anchors for the characters' actions must be introduced and maintained in the discourse too. In the introduction of spatial reference points, the definite article is often appropriate for their initial presentation due to their secondary, backgrounded, and commonly known nature. Speaker's reliance on general world knowledge and inferences based on it are greater in the spatial domain, since it allows for the inference of entity locations without explicit specification (Hickmann, 2003b). Some entities can remain presupposed or not mentioned at all. Usually, inanimate referents are not central to narratives and not all details about locations need explicit mention. The choice of role, form, and predicate type depends on whether children introduce relevant referents in the initial setting or later in the story (Hickmann et al., 1998).

At seven years, late first mentions and the absence of mention are still frequent, but they decrease after this age, while early first mentions increase, becoming most frequent at the adult age (Hickmann, 2003b).

2. Acquisition of definiteness in bilingual children

Keeping in mind all the factors influencing the expression of definiteness in children, I now move on bilingual acquisition in this domain. Compared to monolinguals, fewer studies have investigated the pragmatic adequacy of determiners in bilingual children. The existing results do not provide a clear consensus on when bilinguals develop the ability to use definite and indefinite expressions appropriately or on the potential differences between bilinguals and monolinguals in this aspect (Lindgren et al., 2022).

Serratrice (2007) showed that simultaneous bilingual children (8-year-old English-Italian) can achieve a high degree of language-specific discourse-pragmatic competence in both of their languages. The aim of her study was to compare the referential choices made by simultaneous bilingual and monolingual children in the expression of discourse cohesion in oral narratives in Italian and English. The results of this study show that children who are regularly exposed to two languages from birth can achieve remarkable levels of nativeness in both of their languages. Nonetheless, reference maintenance, especially in the domain of overt pronominal subjects, appears to be more challenging and vulnerable to cross-linguistic influence.

Several other studies have showed that referential choices in narratives by bilingual children are similar to those made by their monolingual peers in both languages (Andreou et al., 2015; Fichman & Altman, 2019; Fichman et al., 2020; Finnstedt, 2013; Topaj, 2010). Serratrice and De Cat (2020) found that bilingual children aged 5-7 years were as knowledgeable about the choice of referential expressions as monolingual peers when their language proficiency in English was controlled for and

they also found that the same cognitive underpinning are in play, without differences between the two groups.

Antoniou et alii (2020) investigated a wide range of pragmatic phenomena (relevance, scalar, contrastive, manner implicatures, novel metaphors, irony and informativeness) and found that school-age French-Dutch bilingual and West-Flanders bilectal children performed on par with their Dutch-speaking monolingual peers, despite lower language proficiency (see Groba & De Houwer, 2018 for an overview on pragmatic abilities in bilinguals). These results imply that at least some pragmatic principles are universal and are not affected by specific language properties (Meir & Novogrodsky, 2021). On the other hand, ways of marking the difference between shared or new information are language-specific, as there are well documented cross-linguistic differences with regard to definiteness marking (see chapter 2).

Zdorenko and Paradis (2012) examined the use of articles in narrative productions in English of four groups of bilingual children (5;0-6;0 years) grouped by family language: two groups with an articleless language, namely Chinese and Urdu/Punjabi/Hindi and two having a language with articles (Spanish, and Arabic). They found that all groups tended to make the same error, specifically, the use of definite articles in indefinite contexts, but that children with family languages lacking articles made more omission errors in all contexts. Both the Spanish-speaking and Arabic-speaking groups exhibited similar patterns for definite and indefinite articles, despite Arabic lacking an indefinite article. According to the authors, this finding would indicate that what is transferred appears to be merely the knowledge of the functional projection D, but not the particular mappings of feature clusters onto morphological forms, i.e., Arabic bilingual children wouldn't transfer the feature [-def]. The authors concluded that bilingual mastery in the use of determiners follows a similar path as monolinguals and in the children's two languages. They argued the mastery is related to developmental issues rather than cross-linguistic influence and that the feature [-def] is inherently more challenging to acquire in both languages. These findings argue against transfer and might reflect the maturation of semantic and/or discourse-pragmatic knowledge.

On the other hand, there is research finding that bilinguals differ from their monolingual peers in the marking of definiteness (Hervé & Serratrice, 2018; Kupisch, 2007; Kupisch & Bernardini, 2007), strengthening the language-specific aspect of referential use. Specifically, lower accuracy on definiteness marking has been linked to the effects of cross-linguistic influences, i.e., the influence of a second language that does not have an article system (Andreou et al., 2020; Chondrogianni et al., 2015; Schwartz & Rovner, 2015). Hervé & Serratrice (2018), studying French-English bilingual children (2;4-3;7) found accelerated progress in English and a slight delay in French in the use of determiners. However, a unidirectional influence from English to French was evident in the significantly higher rate of ungrammatical determiner omissions in plural and generic contexts (areas of divergence in the use of articles between the two languages) compared to singular specific contexts in French. These results suggest that cross-linguistic influence might be due to the children's overall expressive skills in French and that the encoding of definiteness could be delayed in bilingual children due to their lower expressive skills.

In a recent paper, De Cat (2022) made a very important point to better understand the mixed picture resulting from this amount of research. With particular reference to narrative tasks, she stressed that a key challenge in the analysis of bilingual data is to tease apart the source of errors in the use of referential expressions and dissociate grammaticality (e.g., absence/presence of determiners in obligatory contexts) and pragmatic adequacy (e.g., use of the definite article instead of the indefinite article and vice versa).

One possible explanation of the mixed nature of research findings is that bilinguals differ from monolinguals when the languages have different types of referential systems, such as English and Chinese (Chen & Lei, 2013; Chen & Pan, 2009; Jia & Paradis, 2015) or Russian and German (Reichardt, 2014; Topaj, 2010), but that bilinguals speaking languages with relatively similar referential systems, e.g. English and Italian (Serratrice, 2007) or Swedish and English (Finnstedt, 2013), may perform similarly to monolingual age peers. However, despite having similar referential systems, bilinguals might exhibit differing performance in both languages. In a study of narrative

productions of two German–Greek bilingual groups aged 8 to 12, one living in Greece and the other living in Germany, Andreou et al. (2015) observed differences between bilinguals and their monolingual peers in Greek, but not with their monolingual peers in German. Bilingual children living in Germany were more extensively exposed to German than to Greek, while the bilingual group living in Greece had a balanced exposure. This difference in the amount of exposure was reflected in children's referential choices. In fact, despite a generally comparable performance between bilinguals and monolinguals in both languages, most differences between bilinguals and monolinguals were found in the stories in Greek of bilingual children living in Germany. The authors attribute these differences to variations in language exposure, emphasizing the role of input in language acquisition. Torregrossa et alii (2018) argued in favor of dominance in explaining low referential accuracy in bilingual children. In contrast to bilingual children with a predominant experience in the non-target language, those more balanced are likely to have achieved a faster and more efficient proceduralization of grammatical knowledge. Consequently, they can more consistently rely on the syntactic options for reference available in the language.

In the study of bilingual referential skills, particularly regarding the expression of definiteness, the main factors considered causing differences with monolinguals are cross-linguistic influence and factors of individual variability. "Variability is the norm" (Valian, 2020:1) seems to be the acknowledged key point to interpreting the non-monolithic construct of bilingualism; this variability shows itself in cognitive abilities and linguistic experience (De Cat & Unsworth, 2023; Paradis, 2023; Valian, 2015, 2020). The following paragraphs will discuss cross-linguistic influence and factors of individual variability.

3. Cross-linguistic influence in bilingual children

Since the 1970s, there has been debate about the differentiation of linguistic systems in bilingual development. The first proposal came from Swain (1972) and Volterra and Taeschner (1978), among

others, who elaborated on the Unitary Language System (ULS) hypothesis. According to them, bilingual children begin acquisition with a single, undifferentiated linguistic system, like monolinguals, that later separates, typically around the ages of two (Vihman, 1985) or three. This hypothesis was primarily based on the presence of code-mixing in children's utterances. Genesee (1989) questioned this model from both a methodological and an empirical standpoint. According to the author, there is doubt regarding whether code-mixing is a reliable measure of an underlying unitary system. He argued that the presence of code-mixing in bilingual utterances is governed by pragmatic and sociolinguistic factors, which need to be distinguished from grammatical competence, and that the absence of differentiation at the pragmatic level does not necessarily imply a fusion at the grammatical representation level. This view is known as Autonomous Development Hypothesis (ADH). Some years later, Paradis and Genesee (1996) argued that children exhibit very early syntactic differentiation, and even accepting that separation occurs at two years, the question remains open as to whether the systems interact during linguistic development. Their proposal was, therefore, of an autonomous and interdependent development of the two systems (Interdependent Development Hypothesis, IDH). They define interdependence as "the systemic influence of the grammar of one language on the grammar of the other language during acquisition, causing differences in a bilingual's patterns and rates of development in comparison with a monolingual's" (1996:3). They stressed the importance of demonstrating that cross-linguistic effects must be systematic to be considered indicative of influence at the level of language representation.

The separation and interdependence of the two linguistic systems in bilingual children is currently widely accepted in the debate on language development (De Houwer, 1990; Genesee, 1989; Meisel, 1986, 1989; Paradis, 2001; Paradis & Genesee, 1996) and relies on three main assumptions (Meisel, 2007; Serratrice 2013):

1) Bilingual children distinguish between their two languages from an early age.

2) The linguistic development of bilingual children follows the same trajectory as of monolingual children.

3) The grammatical knowledge that bilingual children eventually acquire in each of their languages is qualitatively indistinguishable from that of monolinguals.

Serratrice (2013) noted that ADH in its strictest form, which denies any qualitative distinctions between monolingual and bilingual development, poses some problems, since it is incompatible with the concept of interaction across languages. Furthermore, the evidence regarding ultimate attainment does not strongly support the assertion that simultaneous bilingual children ultimately acquire knowledge qualitatively identical to that of monolinguals. It is crucial to understand that this does not imply that simultaneous bilinguals cannot achieve the same competence as monolinguals. Still, it does suggest that such attainment is not an automatic outcome of early exposure to two languages and underscores the significance of the relative exposure levels to each language. The phenomenon of mutual influence has been called in various ways in the literature: interdependence (Paradis & Genesee, 1996; Meisel, 2001), interference (Bergman, 1976), intrusion (Vihman & McLaughlin, 1982), cross-linguistic transfer (Meisel, 1983) and convergence. The term cross-linguistic influence (CLI henceforth) will be preferred in here since it is the most widely used label in the current research (Serratrice, 2013).

CLI can manifest itself as both a quantitative and a qualitative difference. The former can be observed as acceleration or delay[49] in acquiring certain structures, that is, the impact of language B on language A may be reinforcing (i.e., reducing acquisition times) or slowing down the acquisition of a phenomenon that is also observed in monolingual development. In the case of qualitative differences, a phenomenon is attested in the language A, but it is unattested in monolingual acquisition

[49] In accordance with de Houwer (2021:59), it is acknowledged here that the labels 'accelerated' and 'delayed' are not the most suitable for describing bilingual development, as they entail a monolingual-centered perspective. However, in the absence of well-established substitutes in the field, they will be used for the sake of clarity in exposition. For a critical review of labels used in the field of bilingual acquisition, refer to Kupisch and Rothman (2016) and Surrain and Luk (2017).

of the same language. We define this qualitative difference in terms of transfer when the phenomenon comes to the language A from the language B. These qualitative differences sometimes emerge as overuse or overacceptance of a linguistic property in bilingual children's one language under the influence of their other language (Van Dijk et al., 2021).

Acceleration in bilinguals, especially with Romance-German language combinations, is confirmed by several studies (Bernardini & Schlyter, 2004; Gawlitzek-Maiwald & Tracy, 1996; Patuto et al., 2011). Kupisch (2007) presented the results of a longitudinal study involving four Italian-German bilingual children, two of whom were balanced (Carlotta and Lukas) and two unbalanced (Jan and Marta), each with a different dominant language. Comparing bilinguals with Italian and German monolinguals, in three out of the four cases (Carlotta, Lukas and Marta), the acquisition of determiners in German was faster in the bilingual children than in their monolingual German peers. After the contrastive analysis of the article systems in the two languages, Kupisch showed that the Italian system is more beneficial for acquisition, as it is less complex according to the measures she adopts (2007: 61); research with monolinguals indeed demonstrated that Italian children acquire determiners before German children (Guasti et al., 2004, 2008). Kupisch interprets the results of the three bilingual children as a case of quantitative cross-linguistic influence, where the beneficial language (Italian) supports and accelerates the acquisition of the same structure in the other language (German).

The delay in the bilingual acquisition trajectory emerges in various studies with different language combinations (Patuto et al., 2011; Pérez-Leroux et al., 2017, among others). Some research on object omissions has provided evidence of a delay in the emergence of clitic objects in bilingual children in language combinations where object omission was allowed only in one of the two languages (Hulk & Müller, 2000; Yip & Matthews, 2000). The results demonstrated that in the language where omission was not allowed, the object took longer to emerge compared to monolinguals, due to the language in which omission was possible.

However, Pirvulescu et al. (2012), specifically selected a language pair where object omission is not allowed in either language: English and French. Bilingual children, when compared to monolinguals, omitted the object in French significantly more often than monolinguals. The authors highlighted that the prolonged retention of the object is not due to cross-linguistic influence but to a more general bilingual effect (see also Sorace et al., 2009 for similar conclusions on syntactic-pragmatic phenomena). The delay may, therefore, result from reduced input or dealing with two languages (processing). It therefore becomes necessary to distinguish the general effects of bilingualism from the action of cross-linguistic influence. In a recent meta-analysis, Van Dijik et al. (2021) recommend differentiating these two effects (as in Pirvulescu et al., 2014; Serratrice et al., 2009, 2012) and propose using an appropriate bilingual control group in future studies (as in the case of Kaltsa et al., 2019; Serratrice et al., 2009, 2012; Sorace et al., 2009) or "the introduction of multiple within-experiment conditions that test the same cross-linguistic effects in different ways, and/or the inclusion of matched control-conditions in which only general bilingual effects would be expected (e.g., complete-overlap conditions)" (Van Dijik et al., 2021:923).

Regarding more qualitative differences, numerous studies have reported them in various linguistic combinations. Döpke (1998), for example, observes the presence of V_XP structures in complex verbal constructions in productions in German by bilingual children with English as the other language, where this order is the only possible one. In a more recent study with bilingual French-Dutch children, Strik and Pérez-Leroux (2011) document cases of wh-in-situ questions in Dutch influenced by French. Cases of overuse are reported in Serratrice et alii (2004), where it has been suggested that cross-linguistic influence from English is the reason why, compared to monolingual peers, Italian-English bilingual children overuse overt subject pronouns in a pragmatically suboptimal manner in Italian.

3.1 Conditions for CLI

Given that CLI is a systematic feature of bilingual development, a crucial inquiry in the research on child bilingualism is to define the conditions (specifically, in terms of grammatical domains and exposure) under which this influence can occur.

3.1.1 Linguistic-internal conditions

If interaction between languages is a systematic phenomenon, as asserted by Paradis and Genesee (1996), an account predicting its occurrence conditions has to be devised. At the beginning of the new millennium, Hulk & Müller (2000; Müller & Hulk, 2001) were the first to propose a hypothesis to account for the conditions under which cross-linguistic influence occurs. Their proposal posits that cross-linguistic influence occurs if two conditions are met:

1. There must be partial overlap between the two languages concerning the structure under examination.
2. The structure under examination must be at the syntax-pragmatics interface.

The first condition suggests that in one of the two languages, there must be ambiguity, meaning that in Language A a given construction is analyzable in more than one way, while in Language B evidence is present for only one of the two ways. This condition predicts unidirectionality of CLI, from Language B to Language A. The authors also emphasize that the occurrence of these conditions only makes CLI probable, and thus they are necessary but not sufficient (Hulk & Müller, 2000:229). Serratrice (2013:7) argues that this type of formulation renders the account unfalsifiable, as it is unclear what other factors contribute to making influence more likely. The proposal by Müller and Hulk has been successful, sparking considerable debate and highlighting the need to understand the

factors limiting CLI occurrence in bilingual children. The most well-studied predictors of CLI, which will be considered here, are: surface overlap, language domain, language dominance, and age.

The first condition in Müller and Hulk's formulation is deemed necessary, so conditions of total overlap (both languages behave identically regarding a certain feature) or no overlap (both languages behave completely differently regarding a certain feature) should not lead to any CLI. Many studies have confirmed the prediction and direction (Argyri & Sorace, 2007; Austin, 2007; Haznedar, 2007; Serratrice et al., 2012; Sorace et al., 2009), but many others have found evidence of CLI even in the absence of overlap with different constructions (Nicoladis, 2002, 2012; Nicoladis & Gavrila, 2015; Yip and Matthews, 2000). Results of this nature question the predictions made by the hypothesis of partial surface overlap. Strik and Pérez-Leroux (2011) studied the production of wh-questions in French-Dutch bilinguals (French allows questions with or without inversion and with wh- in situ or fronted, while Dutch only has wh-questions with inversion and with fronted wh-) and found the adoption of wh-in situ questions in Dutch by these children and the absence of subject-verb inversion in wh-fronted questions (both embedded and matrix) in Dutch. This study provides a counterexample to the prediction made by the ambiguity condition, since CLI occurred from the language that provides two options to the one with only one option. The authors propose redefining the concept of surface overlap and suggest an approach based on derivational complexity, where complexity is defined not in terms of the number of structural options but in terms of the number of Merge operations involved in the computation. The prediction of this approach is that the less complex the derivation of a structure, the faster it will be acquired, and this will be the favored structure in both languages. A similar approach has its basis in Jacubowicz's proposal of derivational complexity (see Jacubowicz, 2004, 2005; Jacubowicz & Strik, 2008; Strik, 2009), based on a principle of economy (see also Lightfoot, 1991; Roberts, 2001 for similar proposals). Sorace et alii (2009) also invoke a principle of economy (the so-called "Principle of Avoid Structure") to explain the acceptance by Italian-English bilingual children of bare noun subjects in specific and generic contexts in Italian. In that case, the direction proposed by the surface overlap condition is not satisfied, as CLI occurs

from the ambiguous language (English presents two options: bare noun for generics and definite article for specifics) to the language with only one option (Italian requires the definite article in both cases). The two studies presented here do not falsify the condition proposed by Müller and Hulk but rather complement and integrate it (see Gavarró, 2003, for a similar approach based on a principle of economy).

The second condition posed by the initial formulation of Hulk and Müller (2000) has been tested and refined by subsequent studies. In fact, although numerous studies have confirmed that the syntax-pragmatics interface is particularly vulnerable to cross-linguistic influence (Allen et al., 2008; Allen & Schroder, 2003; Guerriero et al., 2006; Hacohen & Schaeffer, 2007; Paradis & Navarro, 2003; Serratrice et al., 2004. For a similar proposal in adult bilinguals, see Sorace & Filiaci, 2006), many studies have sought evidence for the presence of CLI in other linguistic domains.

In particular, the syntax-semantics interface seems to be vulnerable to cross-linguistic influence. Recall the results of Serratrice et al. (2009) on generics, obtained with acceptability judgments from Italian-English bilingual children, who accepted more bare nouns in generic contexts in Italian than their monolingual peers due to English. Another area lying in the same interface is copula acquisition, investigated in English-Spanish bilinguals by Fernández, Fuertes and Liceras, (2010), Liceras, Fernández Fuertes and Alba de la Fuente (2012). In these two studies, the authors found an acceleration effect of Spanish on the acquisition of the copula in English. The copula in Spanish lexicalizes the distinction between Stage Level predicates (*estar*) and Individual Level predicates (*ser*), resulting in greater lexical transparency. This lexical transparency is for the authors what facilitates acquisition in bilinguals (2 and 3 years old), who showed fewer copula omissions in English compared to their monolingual peers (different results emerged from the study on the copula by Silva Corvalàn & Montanari, 2008).

However, not only interface phenomena constitute areas of particular vulnerability; several studies have found indications of CLI even in more purely morphosyntactic phenomena, such as nominal compounds. Nicoladis (2002) studied cross-linguistic influence in the realization of N+N

compounds in English-French bilinguals. This is a derivational morphology phenomenon, where there is no overlap between the languages, which have two completely different compound structures: head-final in English and head-initial in French. Bilingual children produced head-initial compounds in English 35% of the time, a phenomenon attributed to the influence of French (similar results can be seen in the same author's study on deverbal compounds, Nicoladis, 2003). Foroodi-Nejad and Paradis (2009) also found evidence of CLI in the domain of nominal compounds by studying Persian-English bilinguals. Persian allows both types of compounds with head-final and head-initial structure, resulting in partial overlap with English. Bilingual children produced more ungrammatical compounds in English than monolinguals and a greater number of head-final compounds in Persian. This study not only confirms that CLI also occurs in non-interface phenomena but also that it can be bidirectional. Over the last few years, evidence of CLI has accumulated in structural phenomena in the morphosyntactic domain with various language combinations (Pérez-Leroux et al., 2011; Yip & Matthews, 2000, among others).

3.1.2 External conditions

The conditions described above are language-internal conditions, defined in terms of structural compatibility (ambiguity or derivational complexity) and domain vulnerability. However, external factors that may favor CLI, such as dominance and age, should also be considered.

Regarding dominance, researchers' opinions are not unanimous since the results are mixed. Bilinguals typically have a dominant language and a weaker one (Grosjean, 1982). How to operationalize the dominance of a bilingual child is another question where there is no agreement, and different paths are often followed. Mainly, in the literature, dominance is calculated either based on the input the subject receives (the dominant language is the one in which the subject receives more input, which is usually the societal language), or based on proficiency (the dominant language is the

one in which the child is more proficient)[50]. Regarding cross-linguistic influence (CLI), in a recent article, Serratrice (2022) argues that dominance is a promising predictor. Operationalized as the societal language, Van Dijk et al. (2021) also identified dominance as a significant predictor of CLI. Yip and Matthews (2000) argued that the transfer of wh-in situ questions from Cantonese to English was favored because Cantonese was the dominant language of the participants (calculated formally using MLU as a unit of measure). In the study described above by Serratrice et alii (2009), the authors highlight an effect of the community's language, as English-Italian bilinguals living in England had significantly worse performance than their counterparts raised in Italy. Similarly, Pérez-Leroux et alii (2009) demonstrated that French-speaking children growing up in a bilingual environment significantly differ from monolingual counterparts in rates of production of null objects. Silva-Corvalàn and Montanari (2008) also attribute the presence of CLI found in their study on copula in English-Spanish bilingual children to dominance. However, others are not convinced in the same way that dominance is a good predictor of CLI. Cantone et al. (2008) argue that it is not, as CLI occurs even in balanced subjects and from the weaker language to the stronger one. Indeed, there are studies that have found no effect of dominance in inducing CLI (Nicoladis, 2002; Unsworth, 2012) or unclear effects (Foroodi-Nejad & Paradis, 2009). Van Dijk et alii (2021) argue that, taken together, these results suggest that language dominance, operationalized as societal language, does not predict the presence of cross-linguistic influence, but rather its strength.

Regarding age, the debate is long-standing, and I will refer to the results of the recent meta-analysis by Van Dijk et alii (2021). The authors hypothesize that if CLI is part of being bilingual, it must persist throughout the lifespan, and indeed, their analysis seems to confirm this hypothesis. This is in line with those previous studies that found cross-linguistic influence to remain present in older bilingual children (Argyri & Sorace, 2007; Bosch & Unsworth, 2020; Kaltsa et al., 2019).

[50] See Torregrossa et alii (2021) for an alternative combined way to operationalize dominance in bilingual children.

3.2 Processing accounts and CLI

Many studies have considered the fact that dealing with two languages on a daily basis increases the cognitive load of linguistic processing, and this may also play a role in explaining cross-linguistic influence (Tsimpli et al., 2004 in adult bilinguals, adapted to children in Serratrice et al., 2004; Hervé et al., 2016; Hervé & Serratrice, 2018; Nicoladis, 2006, 2012; Nicoladis et al., 2010; Pérez-Leroux et al., 2009; Serratrice, 2007, 2009). It has been shown how processing issues can be the cause of a general bilingual effect and can explain differences in output between monolinguals and bilinguals. Cross-linguistic influence in bilinguals also seems to find an explanation by resorting to processing issues. In other words, many studies have investigated whether there is an influence of one language on the processing of the other and vice versa (Hartsuiker et al., 2016).

While many studies have focused on the representation or processing of words (Dijkstra & Van Heuven, 2002), others have addressed syntactic representations, which will be the focus of this section, as our interest lies in the nominal phrase. Recent research on syntactic representations in bilingual adults and children has benefited from the adoption of the structural priming paradigm.

Structural (or syntactic) priming is known as the propensity for individuals to repeat the same or a similar syntactic structure they have just heard someone else employ or used themselves (Bock, 1986). The rationale behind priming is that the processing of a syntactic structure facilitates its subsequent use (Bock, 1986; Serratrice, 2016). Syntactic priming can work not only within languages in monolingual speakers (within-language priming), but across languages too (between-language priming) in bilingual individuals (Serratrice, 2016). That means, for example, that listening to a passive sentence in Spanish would facilitate the subsequent use of the passive in producing an English sentence by a bilingual Spanish-English speaker.

The effects of cross-linguistic structural priming are consistent with two different ways of organizing syntactic information in bilingual individuals: syntactic representations are shared between the two languages (shared-syntax accounts), or they can be separate but still interacting

(separate-but interacting syntax accounts). These accounts are complex, and their main features are summarized here, prioritizing synthesis over exhaustiveness (for further details, readers are referred directly to the works cited in the text). These psycholinguistic proposals are presented here, albeit briefly, as they offer a promising explanation of cross-linguistic influence based on processing and as they can be linked to speech production models. It should be acknowledged that these processing accounts have been primarily developed through experiments involving adults; relatively little research has focused on children. The current two main accounts are:

Shared syntax account: there are two versions of this model, one lexicalist (i.e., where the transfer process is lexically driven) and one non-lexicalist. Hartsuiker et alii's (2004) lexicalist shared-syntax model posited an integrated level of lexical and syntactic representations in bilingual mind, whereby syntactic representations are shared between languages whenever possible. This model is a multilingual expansion of Pickering and Branigan's (1998) framework concerning the lexicosyntax stratum in sentence production (lemma stratum; see also Levelt, 1999). Hartsuiker et alii's model incorporates lexical representations for words in each language. These lemma nodes are connected to conceptual nodes (capturing meaning), to combinatorial nodes (both containing syntactic information) and to language nodes (indicating language membership). It is assumed that the conceptual and combinatorial nodes are shared across the two languages. To clarify the model, Unsworth (2023: 3-4) provides a picture representing the production of possessive structures in Dutch-English bilingual children (Figure 2A). When bilingual children hear prenominal possessives in Dutch (e.g., *de non haar ei*,'the nun's egg'), this activates the combinatorial node associated with that structure, and due to the residual activation from that operation, the same structure becomes subsequently more available when the same speaker needs to produce a possessive, even in the other language (English, 'the nurse's horse' instead of 'the horse of the nurse'). There would therefore be a single combinatorial node connected to the prenominal possessive structure, and each noun in every language the speaker knows would be linked to that structure. The same applies to the other structure (Hartsuiker et al., 2016).

There is also a non-lexicalist version (i.e., implicit learning), akin to the one proposed by Chang, Dell, and Bock (2006). These frameworks posit that syntax operates independently of specific lexical items, suggesting that in multilingual individuals, syntax could abstract away from language distinctions, provided that the structures of the languages involved exhibit similarities.

Separate-but-interacting-syntax accounts: another perspective on syntactic organization in bilingual individuals posits that they maintain distinct syntactic representations for each of their languages, yet these representations interact with one another. For example, De Bot (1992) introduced a bilingual adaptation of Levelt's (1989) model, which distinguished a conceptualizer, a formulator, and an articulator, along with a shared lexicon between production and comprehension. De Bot suggested that processing in both languages overlaps at the conceptual and lexical levels, but separate and interacting formulators exist for each language. The degree of interaction depends on factors such as the etymological similarity between the languages (with greater linguistic distance resulting in diminished cross-linguistic influence) and L2 proficiency (higher proficiency in L2 facilitating clearer separation between the two languages and thus reducing cross-linguistic influences). Kantola and Van Gompel (2011) also proposed a lexicalist model with separate but interacting syntax, known as connected-syntax account. Aligned with the model proposed by Hartsuiker et alii (2004), this approach posits that structural priming arises from residual activation of combinatorial nodes and their connections. Lemmas from both languages are housed within a unified lexicon, yet unlike Hartsuiker et alii's model, combinatorial nodes are represented separately for each language. In Unsworth (2023)'s example (Figure 2 B), prenominal and postnominal possessives are duplicated, existing in both English and Dutch. The strength of those connections is determined by the degree of resemblance between the structures. Activation of these connections can also be enhanced through structural priming.

Regardless of the approach, it is important to stress that most of the studies showed that lexical activation operates without strict language selection. When bilingual individuals produce words in

one language, it triggers activation of related words in the other language. Those findings suggest that when bilinguals engage a syntactic structure in one language, they activate a syntactic representation that lacks language specification and is common to both languages, if the grammatical structure of this configuration is similar across both languages (Kantola & Van Gompel, 2011).

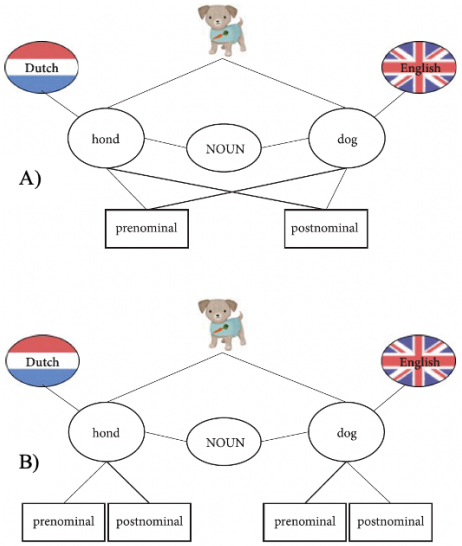


Figure 2 Models for the representation of Dutch and English possessive constructions in English-Dutch bilinguals (lemma stratum). A) shared-syntax account, B) connected-syntax account. From Unsworth (2023

These models provide a useful psycholinguistic description of the phenomenon of our interest: cross-linguistic influence. Serratrice and colleagues (Hervé et al., 2016; Serratrice, 2016, 2022) have suggested that one way to think of CLI is as between-language priming, which is the outcome of previous exposure to one language and of common syntactic representations between languages.

However, to date, there is limited research on crosslinguistic priming in younger bilingual children compared to adult L2 learners. Adult L2 speakers and simultaneous bilingual children have in common that their output is often indistinguishable from that of monolinguals, but they differ fundamentally in that the former are already fluent in their L1 and usually quite so in L2, while the latter must make structural sense of input from two languages. Children who are bilingual often differ in the possibilities to come across potentially similar syntactic constructions in both of their

languages, due to varying exposure levels and diverse situations. They thus require more time to establish equivalency across languages and it may take longer for structures to become shared in the event of the asynchronous exposure across languages (Serratrice, 2022). In order to account for developmental issues, Hartsuiker and Bernolet (2017) introduced a developmental framework of shared-syntax account. According to them, in the initial phase of second language (L2) acquisition, L2 vocabulary is stored independently from structural connections. Subsequently, bilingual individuals establish connections between L2 words and structural representations, yet these structural representations remain language-specific and fully tied to lexicon. Eventually, structural representations in both L2 and L1 become shared across languages, culminating in the situation described by the shared-syntax model.

The existing studies on bilingual children are few and focus on the priming of transitives in Spanish-English bilingual children (Gámez & Vasilyeva, 2020; Vasilyeva et al., 2010), on the priming of adjectival constructions in the same language combination (Hsin et al., 2013), on the priming of ditransitive constructions in Norwegian-English bilingual Children (Wolleb at al. 2018) and on the priming of possessive constructions in Dutch-English bilingual children (Unsworth, 2023). Overall, these studies demonstrated that CLI could be primed, suggesting a common representation of abstract syntactic structures in bilingual children, like in bilingual adults.

A factor, extensively debated in language acquisition (see Gülzow & Gagarina, 2007), that has also been reconsidered in the study of cross-linguistic priming and that seems to play a decisive role is frequency. For bilinguals, the frequency of a specific structure may be influenced by both languages, particularly when the structure is shared between them. Shared representations have been observed to "inherit" the structural frequency of the language that is not currently in use (see Runnqvist et al., 2013). The most frequent structures become more entrenched representationally and as such more easily accessible during processing. In this account the notion of "inherited frequency" is in use (Unsworth, 2023). Hervé and Serratrice (2018), in an exploratory corpus-based study on the development of determiners in two English-French bilingual children, found that instances of CLI

(such as the omission of articles in obligatory contexts in French under the influence of English) were lexically specific and occurred with high-frequency words (e.g 'chocolate' and *chocolat*). Their findings suggest that the degree of activation of language-specific form-function associations may also depend, to some extent, on lexical factors. Therefore, CLI may receive a lexical boost when a high-frequency competing structure is simultaneously activated.

Two additional factors, at least, have been extensively considered in the study of cross-linguistic priming: the similarity between structures and the proficiency of the speaker.

The first, although crucial, has not yet been fully clarified, and further studies are needed. In other words, it is not clear to what extent competing structures need to be similar for their representations to be connected or shared. As seen above, structural similarity is posited as a necessary condition by several accounts. De Bot (1992) argues that the more similar the two structures are, the easier it is for common representations to exist, leading to a higher probability of CLI (the more distant the languages, the less CLI. Hartsuiker et alii (2016) also argue that structural similarity is important for CLI. They address this issue for adults and argue that for priming to occur, morphological or pragmatic differences do not seem to be important, but instead, identity in word order seems necessary. Syntactic structures with different constituent orders between the two languages do not seem to lead to priming and thus seem to have separate representations (Hartsuiker et al., 2016). Regarding bilingual children, the need for word order identity is confirmed by some studies and contradicted by others, and therefore needs to be further investigated (see Serratrice, 2022). Moreover, Unsworth, in line with Hervé and Serratrice (2018), argues in favor of a lexically-driven shared-syntax account, according to which syntactic structures are shared only between items which are translation equivalents of each other. Even in this case, there is no agreement, and the issue seems still under-investigated.

The second factor to consider is proficiency. Again, the results of studies are quite mixed. According to De Bot (1992), the more proficient an individual is, the easier it is to separate representations and the less likely CLI is, thus his model predicts that the probability of CLI decreases

with increasing proficiency. Different results come from Hartsuiker and Pickering (2013) and Hartsuiker and Bernolet (2017), according to which priming would be higher for individuals with higher proficiency. It should be noted, however, that these studies are often conducted on highly proficient adults (typically university students). Unsworth's (2023) results seem in line with previous studies by Bernolet et alii (2013) and Hartsuiker and Bernolet (2017), in which cross-language priming was influenced by proficiency, as also noted by Wolleb et al. (2018). Unsworth's analysis indicated that proficiency effects, when observed, were predominantly influenced by children's proficiency in their heritage language (i.e., English or Spanish) rather than societal language (Dutch). Specifically, in cross-language priming, higher proficiency in the heritage language correlated with a stronger priming effect. Put differently, greater proficiency in the heritage language was linked to increased utilization of the structure corresponding to the most prevalent or exclusive structure in that language overall (Unsworth, 2023).

3.2.1 Shared Syntax and CLI in the speech-production model

The results obtained from the aforementioned studies on bilingual children are, in principle, compatible with both shared syntax and connected syntax accounts (Unsworth, 2023:28). In this work, we are not able to determine which of the two models is more explanatory, nor is this the objective. The merit of these models lies in their consideration of processing issues in explaining cross-linguistic influence during production. We will refer to this family of models as processing accounts (following Hervé and Serratrice, 2018). This type of psycholinguistic models, concerning production, is linked to speech production model, which has already been presented in 3.1.2 (based on Levelt, 1989).

Of particular interest for understanding CLI in bilingual children is the influential proposal coming from Nicoladis (2006, 2012)[51]. The author defines CLI as an epiphenomenon of speech

[51] Nicoladis (2006, 2012) proposal is based on models of connected-syntax (Kantola & Van Gompel, 2011).

production. Based on various previous models (Dell et al., 1999; Dell at al., 2000; Ferreira & Dell, 2000; Costa, 2004), Nicoladis proposes a speech production model with the three levels (or strata) already mentioned (see 3.1.2): the concept level, the lemma level, and the phonetic articulation level. At the concept level, the speaker conceives the message they want to convey. In the second phase, the speaker chooses the specific words and the syntactic structures in which the words will appear. At this level, the language of words and the syntactic structure should be chosen. If there are different syntactic structures (shared or connected), that can convey the same message, there is competition between these structures, which can sometimes result in errors. The last stage is the choice of the phonological form and the actual articulation of the message.

Cross-linguistic influence, according to Nicoladis' proposal, occurs at the lemma level. At any point where the speaker has to make a choice, it is likely that two options in competition are activated, which in bilinguals can come from both languages. At the lemma level, the bilingual child has to choose words from the appropriate language and also from the appropriate syntactic structure. Nicoladis suggests a processing explanation for word order reversals in French–English bilingual children, specifically in Adj + N and N + Adj structures. When a word (e.g., 'apple') and an adjective (e.g., 'green') are activated in English, this triggers the associated English Adj + N word order. Simultaneously, for a French-English bilingual, the translation equivalents ('*pomme*' and '*vert*') will also be co-activated, though to a lesser extent, along with the French N + Adj word order. The co-activation of these structures may result in cross-linguistic influence when the competition is dominated by the structure of the non-target language, such as the French word order when the speaker is employing English words in an English context. A prediction from this processing account is that CLI would be enhanced through syntactic priming following prior exposure and/or use of a specific structure.

In summary, the processing account posits the potential for bidirectional cross-linguistic influence in any language pairing, stemming from the interaction between the two competing linguistic systems. The extent of co-activation would be influenced by factors such as the frequency of the target structure

in the input. Hervé and Serratrice (2018) emphasize that accounts of partial overlap and economy fail to explain all instances of cross-linguistic influence (CLI) in their study (on determiners) in bilingual production. Instead, all instances seem consistent with a processing account linked to the speech production model. Nicoladis' model effectively explains past occurrences of CLI, both in the presence or absence of overlap and beyond the interface condition (Nicoladis, 2002; Yip & Matthews, 2000; Nicoladis et al., 2010). This model suggests that cross-linguistic influence functions as a type of speech error. In this speech production model, effects from overlap, ambiguity, and/or dominance are not excluded, but they are integrated into a single predictive model (Nicoladis, 2006).

4. Factors of individual variability

In the previous section, we explored cross-linguistic influence. In this section, we will briefly examine the factors that serve as determinants of the individual variability characterizing bilingual development. Following Paradis (2011), variables impacting the rates of language acquisition, which differ among individuals, can be classified as either internal or external to the learner. Child-internal factors encompass processing or cognitive factors, and chronological age. Child external factors are mainly factors that determine the quantity and the quality of the input the child receives in the target language.

It has been seen in 3.1 that chronological age, and so cognitive development, plays a significant role in mastering the multifunctionality of articles and, more in general, referential abilities in monolinguals. However, we have also shown that bilingual and monolingual rates of acquisition can be different. In addition to chronological age in bilinguals the age of onset needs to be considered. Age of onset is related to country-specific educational practices because it often coincides with the age at which children tend to start daycare, where in general the country's majority language is spoken (3 years for Italy) (Schulz & Grimm, 2019). Age of onset can distinguish between two groups of bilinguals: simultaneous (2L1) (the exposure to more than one language starts in infancy, up to about

age 2), successive (eL2) (the exposure to other languages starts in early childhood, up to about age 6) (De Houwer, 2021:1)[52]. Schulz and Grimm (2019), in alignment with Tsimpli (2014), argue that to explain the differences in acquisition rates among monolinguals and the two groups of bilinguals, it's not enough to consider only the age of first exposure; the timing of acquisition of phenomena in L1 also needs to be taken into account. In their study, they distinguish between early-acquired phenomena (before the age of 5), late-acquired phenomena (around 5 years), and very late-acquired phenomena (after 6 years). They analyzed six morphosyntactic phenomena in German, involving monolingual, simultaneous bilingual, and sequential bilingual children. The main findings indicate that simultaneous bilingual children have an advantage over their early L2 peers in early-acquired phenomena, which diminishes over time. In late-acquired phenomena, simultaneous bilingual and early L2 children show no significant differences. Additionally, simultaneous bilingual children perform similarly to monolingual children in early-acquired phenomena but face a disadvantage in late-acquired phenomena, with the degree of delay decreasing over time.

Chronological age is strictly related to cognitive maturity. Regarding cognition, a central theme in the bilingualism debate is the so-called “bilingual advantage”. Much research has focused on the effect of bilingualism on domain-general cognitive skills and, in particular, executive functions (henceforth EFs) (Baum & Titone, 2014; Luk, 2022). However, when it comes to children, it is less straightforward to demonstrate the existence of a cognitive advantage in bilinguals and results are mixed (see Valian, 2015, for a discussion on “bilingual advantage” in children). Recent studies highlight the influence of working memory capacity (WM) and EF skills on reference production variation in bilingual children (Hendriks, 2016; Serratrice & De Cat, 2020; Torregrossa, 2017; Torregrossa et al., 2021). WM is associated with storing information about a discourse referent, while EFs are linked to retrieving and updating this information when the referent is mentioned again. In

[52] A third group can be mentioned: children who grew up monolingually throughout infancy and early childhood (up to about age 11) may start attending school in a new second language (L2) that differs from the one people were talking to them before (their L1). These children in middle childhood are growing up in a Second Language Acquisition (SLA) setting (De Houwer, 2021:1, see also Gottardo et al., 2023).

their study in 2017, De Cat et alii showed that bilingual children exhibit better performance in the inhibition function, which seems to underlie all EFs, and this has been interpreted as evidence of the existence of a bilingual advantage (see also Yurtsever at alii, 2023). In 2020, however, De Cat and Serratrice studied the relationship between EFs and the adequacy of referential expressions in bilingual children and did not find evidence of this advantage. It should be noted that the analyzed sample was heterogeneous and not controlled for immigration status, linguistic distance nor interactional contexts. In fact, it has been shown (Valian, 2015, and De Cat et al., 2017) that executive functions have a positive relationship with some external factors such as socioeconomic status and involvement in challenging activities (multimedia, music, etc.). What De Cat et alii (2017) argued is that a threshold exists, determined by external and personal factors, beyond which the advantage does not manifests itself.

From that, there emerges an interrelation between cognitive factors and linguistic experience (Torregrossa et al., 2018). For what concerns referential expressions, like definite and indefinite NPs, Torregrossa et alii (2021)'s findings underscore the substantial impact of language experience on shaping reference patterns in bilingual children (Albanian-Greek). In line with the study's hypotheses, referential accuracy varied based on the dominance (defined both in terms of proficiency and exposure) of the children in either Greek or the non-target language. The preference for overspecified full NPs in their experiment was associated with reduced levels of language exposure in the target language.

Exposure essentially concerns the quantity of the input the bilingual child receives. A child's language exposure may differ based on the overall duration or variations in time at school, in the community, and at home. Input also varies based on more qualitative aspects: differences in experiences with native-speaker input, exposure to rich and complex input through activities such as reading, and interactions with interlocutors whose interactive styles promote language development (Paradis, 2011). Researchers commonly aggregate and quantify specific elements of children's language input to comprehend the influence of these diverse experiences on language outcomes and

operationalize them. This often involves reducing them into broader variables (usually extracted from parental questionnaires; see Paradis, 2017): namely language richness to gauge input quality and the amount of language exposure as a measure of input quantity (Unsworth, 2019)[53]. Many researchers have highlighted the significance of both the quality and quantity of input in influencing the acquisition of linguistic phenomena (Kupisch & Rothman, 2016; Treffers-Daller et al., 2007; Torregrossa et al., 2018; Tsimpli, 2014).

The quantity of input is determined not only by the contexts in which exposure to the language occurs but also by the number of individuals serving as sources of input in that language. In addition to caregivers, various studies have highlighted the importance of the entire social network in which the child is involved, both domestic and not (Chini & Andorno, 2018; Unsworth, 2016). Interactions with siblings (Biazzi, 2018; Macleroy, 2022), and intergenerational interactions can all contribute to language exposure (Biazzi, 2018; Quay and Montanari, 2016;). Sociocultural elements, such as intergenerational parental discourse strategies, the involvement of other family members, parents' attitudes towards the languages (Torregrossa & Carbonara, 2023) contribute to the linguistic environment. This highlights the importance of considering the dynamics of the social context when studying bilingual individuals (Luk & Grundy, 2023; Titone & Tiv, 2023;). Recent research suggests that also social engagement, especially with peers, plays a crucial role in early language learning (Unsworth, 2016). Peers, including friends and classmates, are potential sources of language input, although limited research is available on this aspect. Furthermore, studies suggest that output, or a child's active use of language, can be a significant predictor of language development (Bohman et al., 2010). Actively using a language through output would engage learners in a way that passive exposure through input does not.

As for the quality of input, a significant aspect should be stressed: children from newcomer (immigrant and refugee) families may not receive proficient input in the societal language at home,

[53] The ways in which input characteristics are operationalized vary widely from researcher to researcher, also influenced by the use of different parental questionnaires in terms of structure and objectives (see Kašćelan et al., 2021; Tomić et al., 2023; Unsworth, 2019).

especially when their parents are also in the process of learning the language (Meluzzi et al., 2018:249-251; Paradis & Navarro, 2003). This is due to the presence of non-native or attrited speech (Mehotcheva & Köpke, 2019) in their linguistic environment, reflecting contact-modified input (*contra* Hauser-Grüdl et al., 2010).

After the 1970s, many languages came into contact with Italian as a result of new waves of immigration (see 6.2.1 for a description of plurilingualism in Italy). This led to the arrival of many new adults in the country, of whom the children who participated in the present study are the offspring or grandchildren. Italian learned by adults as a second language has been the subject of various studies by Italian second language acquisition research (see, among others, the volume by Giacalone Ramat, 2017).

The acquisition of articles, focus of interest in this work, is notoriously one of the most problematic areas for learners of Italian as a second language: the development of this linguistic category is indeed very slow in the early stages of acquisition and is subject to fossilization at advanced levels (Chiapedi, 2010). The definite article seems to be especially challenging, above all for adult learners who speaks a first language where there is no overt element comparable to the Italian definite article (as for Chinese speakers; see Andorno; 2010, 2011; Arcodia & Basciano 2020; Chiapedi, 2010; Martari, 2017). This leads to prolonged phases of omission, which can also become fossilized. It is a different case for learners whose L1 and Italian have similar article systems (for Spanish-speaking learners of Italian, see Bailini, 2016:198-201; Calvi et al., 2011; González Luna & Sagi-Vela, 2020:273; for Arabic-speaking learners of Italian, see Abi Aad, 2006; Della Puppa, 2007; Mion, 2020:187; Martari, 2021). In general, the definite article appears later than the indefinite article in adult Italian learners' grammar. Children who hear Italian at home from non-native Italian-speaking parents may be exposed to input with numerous omissions of the definite article. Moreover, L2 adult learners of Italian show difficulties in the domain of morphology, particularly in the agreement between determiners and nouns in gender and number (Giacalone Ramat, 2017).

The issue of input quality is also related to children's parental background, which socio-economic status or educational background parents have (De Cat, 2021). Socio-economic status within the family, frequently assessed by maternal education levels, consistently predicts language development, especially in terms of vocabulary in monolinguals and bilinguals (Hoff, 2003, 2006). Regarding input quality, Jia et alii (2007) measured the "richness" of the English environment, considering factors like native-speaker friends, reading books in English, and engaging in other media in English, highlighting the importance of both native-speaker and rich input (see Sun & Yin, 2020 for input from multimedia in bilinguals). Scheele et alii (2010) explored enriching home language activities in Moroccan-Dutch and Turkish-Dutch families, revealing significant correlations between these activities in L2 and L2 vocabulary outcomes, while Listanti et alii (2023) show that literacy exposure in the heritage language at home enhances children's theory-of-mind abilities in Greek-Italian bilingual children.

In summary, bilingual development is intricately shaped by internal factors such as age and cognitive processing, as well as external factors like input frequency and quality. A nuanced understanding of these dynamics is crucial for unraveling the complexities of bilingual language acquisition.

5. Research Questions and Predictions

Based on the literature presented in the previous pages (chapters 1, 2, and 3), three research questions were formulated, with the corresponding expectations, which are reported below:

RQ1. Do bilinguals and monolinguals distribute articles differently according to identifiability?

Recall that the children are 4-8 years old. Based on the facts that at that age, the cognitive skills underlying pragmatic abilities for reference are already in place (Rozendaal, 2008) and that the

referential systems of Italian and the involved family languages differ to a varying extent, I expect that all children have already developed sensitivity to the distinction of identifiability, but at the same time, the distribution of determiners would be more similar to the one found in monolingual for those individuals whose family language more closely resembles Italian with respect to its referential system. Considering that children were administered a narrative task, cognitively more demanding than spontaneous conversations, it is expected that some errors may emerge. It is expected that the most common error in all groups would be the use of the definite article in non-identifiable contexts (egocentric error).

RQ2. Does the family language influence the societal language in the acquisition of definiteness?

The expectation is that Cross-Linguistic Influence (CLI) from the family language to Italian would manifest in the domain of determiners if there was partial overlap between the two languages:

- for the Spanish-speaking group, a visible effect of the family language on Italian is not expected, given the total convergence of the article systems in the two languages, with respect to the phenomena under investigation. The only area of divergence where the effects of Spanish on Italian can be expected is the co-occurrence of the article and the demonstrative in the same NP.

- As for the Arabic-speaking group, the article systems of spoken varieties of Arabic and Italian substantially converge, with respect to the phenomena under investigation. Areas where the effects of the family language might be evident include the co-occurrence of the article and the demonstrative in the same NP in anaphoric contexts. It is also expected that there would be less extensive use of bare nouns by Italian-Arabic bilingual children compared to monolingual peers, in contexts where Arabic prefers the use of the definite article while Italian allows bare NPs: mass nouns in PPs and bare plurals in non-identifiable contexts.

An effect of Chinese on Italian is expected in two forms; where Italian would use the definite article, Chinese would tend to :

- overuse demonstratives in an anaphoric contexts,
- omitt the definite article in identifiable contexts,

As for the canonical word order, it is the same both in Italian and Chinese: Subject-Verb-Object (SVO), with the subject typically being topical and definite. There is also substantial convergence between the two languages in presentative sentences, which feature a post-verbal subject in both languages (with '*c'è + NP*' in Italian and '*you + Cl + NP*' in Chinese, meaning "there is'" in English; see 2.4.3). Regarding the object, in Italian and Chinese, it is normally post-verbal. However, in Italian, its position remains unchanged whether accompanied by the definite or indefinite article; whereas in Chinese, to receive a definite reading, it must move before the verb; otherwise, it will tend to raise an indefinite reading (see 2.4.3). From this, one might expect that cross-linguistic influence manifests at the level of the object, and that a Chinese-Italian bilingual speaker, when speaking Italian, would tend to place the object in pre-verbal position when they want to signal its identifiability to the listener.

RQ3. Is there an effect of child-external and internal factors on the acquisition of definiteness? Given the interplay between morphosyntax and pragmatics, it is expected that children with greater dominance in Italian morphosyntax would also have higher levels of accuracy at the discourse-pragmatic level (fewer substitution errors) and less article-omissions. It is also expected that greater quantity and better quality of exposure would be linked to more target-like use of articles (fewer omission or substitution errors).

4. Method

The study presented here examines narratives produced by bilingual children to investigate the expression of definiteness. In this chapter, the methods employed to collect and analyze data are outlined. Firstly, the research questions are presented, then a comprehensive description of the participants is provided, highlighting their linguistic backgrounds. The following section outlines the materials and tools utilized to collect data, focusing on their suitability for examining the linguistic phenomena of interest in our multilingual setting. Then, the procedures employed in data collection and data analysis are described: linguistic measures, child external measures, and NP coding.

Three family languages paired with Italian were taken into consideration: Spanish, Arabic, and Chinese. These languages have been purposefully selected for their diverse typological characteristics, as delineated in chapter 2, and their widespread presence in the Italian educational landscape.

1. Participants

The children included in the study[54] were divided into a control group (monolingual children) and three experimental groups (bilingual children with three different family languages). All children were recruited from two cities in Northern Italy, specifically from two schools in Bologna (namely I.C 5 and I.C. 11) and an educational center in the city of Pavia (*Scuola della Pace*, Community of St. Egidio, Pavia). The selection criteria were:

- age between 4 and 7 years,
- (only for the experimental group) being active bilinguals in both languages, that is, Italian (societal language, ITA) and one of these languages: Spanish, Arabic or Chinese (family language, FL).

[54] This research project was approved by the Institutional Review Board of the University of Bologna. See Prot. n. 283813 4th November 2021.

- (for the control group) be monolingual Italian speakers.

Educators and teachers were asked to give the researchers a list of children who met the selection criteria and could be included in the study. The families of these children were given privacy and informed consent forms. 21 (41%) of the 51 bilingual children on the list were excluded from the research project after data collection because they did not produce any narratives. Many possible reasons can account for this: contrary to the educators and teachers' expectations, children were not active bilinguals in the family language and were unable to have even a basic interaction with native speakers, or other factors may have inhibited the production, such as tiredness, unfamiliarity with researchers or tasks etc. Out of these 21 children, 4 were Spanish-speaking, 14 Arabic-speaking, and 3 Chinese-speaking.

41 children were included in the present study, all between the ages of 4 and 7 years except for two children who had just turned 8 at the time of collection. Most of the children in the experimental group were either born in Italy or arrived by the age of three (28 out of 30), except for two who arrived at the ages of 5 and 6. With the exception of these two children, all were exposed to Italian by the age of three, when they had access to formal education in the kindergarten. The control group consisted of 11 monolingual Italian children, with both parents being Italian (only one child had a non-Italian father, who did not reside with the child). The experimental group consisted of 10 children with Spanish as their family language (both parents originating from the same country: Peru, Santo Domingo or Ecuador), 10 children with Arabic as their family language (both parents from the same country: Morocco, Tunisia, Egypt or Syria), and 10 children who spoke Chinese as their family language (both parents from the People's Republic of China). For none of the children teachers

reported history of speech or language impairment[55]. None of the children were taught Italian as a second language at school[56]. Age statistics by language group are reported in Table 1:

Family language	n	Range	Median	Mean (sd)
ITA	11	4;0 to 7;6	6;4	6;35.6 (1;07.9)
CHI	10	4;5 to 8;0	6;9	6;75.8 (0;93.6)
SPA	10	4;1 to 7;0	7;0	6;45.8 (1;03.8)
ARA	10	4;9 to 8;3	6;8	6;77.5 (1;03.8)

Table 1 Overview of age by language group with number of children per group, age range, age median and mean (standard deviation)

A Wilcoxon test was run to assess the significance of the age differences between the groups (Ita-Chi: $p = 0.34$; Ita-Spa: $p = 0.28$; Ita-Ara: $p = 0.37$; Chi-Spa: $p = 0.64$; Chi-Ara: $p = 1$; Spa-Ara: $p = 0.72$). The p-values ranged between 0.2 and 0.3 in the comparisons between the monolingual group and each bilingual group, whereas they exceeded 0.6 in the comparisons between bilingual groups. The results indicate that there are no statistically significant differences in means among the groups.

2. Materials

The collected material consists of:

- a collection of narrative productions in Italian and in the family language,
- the results of a working memory test,

[55] It must be said that in the Italian school system, the Ministry for Education recommends that the diagnosis of learning and language disorders are conducted only after the second year of primary school (Consensus Conference, 2007). As all participants fall within the period preceding that recommended for diagnosis, for this study we are not in possession of clinical reports on the children.
[56] All children were also exposed to English as a foreign language at school, but the influence of this factor was not considered in the present study.

- the socio-biographical information emerging from questionnaires for families and teachers/educators.

Collection of narratives: for the collection of narrative productions the MAIN (Multilingual Assessment Instrument for Narratives) developed by Gagarina et alii (2019a) was used. The instrument has been designed in 2012 (Gagarina et al., 2012) to assess the narrative skills of children aged 3 to 10 years and was later revised and expanded to include adolescents and adults. In this study the revised versions (Gagarina et al., 2019a) were followed.

The instrument is made up of four parallel picture-based stories (Cat, Dog, Baby Birds, and Baby Goats), each with six pictures and a script for the story. The linguistic and cognitive demands of the four stories are comparable, and their cultural appropriateness has been controlled. As a result, they allow assessing bilingual children in both languages and enable comparisons to be made between children from various linguistic and cultural backgrounds. Three elicitation procedures—telling, retelling, and model story—may be used to evaluate children's narrative skills. Only telling and retelling tasks were used in the present study.

The production can then be examined on both a microstructural and macrostructural level. The macrostructure assesses the higher-order narrative structure, which is thought to express universal structures (cognitive schemata) and can therefore be considered language-independent. The microstructure concentrates on elements that are connected to language-specific characteristics, such as the number and complexity of words and sentences, and lexical diversity (Lindgren et al., 2023)[57].

For the purposes of this study, Italian (Levorato & Roch, 2020), Spanish (Ezeizabarrena & Garcìa del Real, 2020), Mandarin Chinese (Luo et al., 2020) and Lebanese Arabic (Gagarina et al. 2019b) revised versions of the MAIN protocol were used. For Arabic, the story script in Lebanese was adapted by native speakers to the regional varieties involved in the study (see Appendix B)[58]. All

[57] The validity of MAIN as an assessing tool was evaluated in the study by Lautenschläger et alii (2021).
[58] At the time of administration, only the Lebanese version was available for Arabic-speaking children.

the stories in the instrument were used: Dog and Baby Goats for Italian, while Cat and Baby Birds for the family language.

Working Memory Assessment: to assess working memory in children, the "Tieni a Mente" test from the FE-PS 2-6 battery was used (Usai et al., 2017). During the assessment the child is shown some pictures belonging to one of the following five categories: animals (dog, fish, cat, mouse), clothing (socks, T-shirt, skirt, shoes), transportation (train, bicycle, motorcycle, car), fruit (banana, pear, strawberry, apple) and sky (moon, sun, star, cloud). The test consists of six trials. The pictures are shown in series of six. Before each trial, the child is asked to pay special attention to one or two designated categories. During the presentation of each series the child must name each picture. At the end of the series, the child must remember the last item of each designated category. The number of categories to be remembered increases from one in the first three trials, to two in the remaining three trials. In order to lighten the memory load, the bottom of the paper displayed small boxes showing the category to be remembered. The task requires continuously updating the information to be retained by the child and managing the interference generated by pictures of other categories. The normative data of the test were collected on a sample of typically developing children aged 48 to 78 months. From the performance recorded in the normative sample, changes by age are linear and gradual: at age 4, children are able to achieve an overall score of 3/9, while at age 6 they are able to recall 4 items out of 9. In the present study the test was administered to children aged over 78 months, and they indeed showed scores in line with the age changes reported by normative data[59].

Questionnaire: for the collection of biographical data and information on language use, parents and teachers were each asked to fill out a questionnaire (see Appendix B). Both the parental and the teacher questionnaires were created specifically for the present study on the model of questionnaires

[59] To assess whether the participant group in this study aligned with the normative sample, a Pearson correlation was conducted between the children's age and their scores in the memory test. Age and Working Memory were found to be positively correlated, $r(39) = 0.61$, $p < 0.001$.

available in Italian (INVALSI, 2009; Roch et al. 2012; Contento et al. 2013; Chini & Andorno, 2018; Fiorentini & Gianollo, 2018).

Teachers/educators were asked to distribute the questionnaire to the families and subsequently return it to the researchers. Many of the questions were shared between the two questionnaires. Most of the information was gathered from the families. Teacher questionnaires were used only when essential information was missing from the parental questionnaires. All questionnaires were administered in Italian, and families had the option to receive clarification in their native language from one of the native speakers involved in the study (who had previously undergone training on the questionnaire completion procedure).

The questionnaire included the following sections:

A. **Biographical and family data**: it gathers information on the child's date and place of birth, household composition, parents'occupation, and their level of education.
B. **Schooling**: it collects information on the year of entry to kindergarten and school attendance.
C. **Language exposure**: it gathers data concerning the exposure to Italian language both at home and in external settings, the sources of the input (including individuals, media, and books), as well as self-assessed proficiency levels in Italian for both the mother and father.
D. **Home learning materials and activities**: it collects data on the frequency and nature of home-based literacy activities in which the child participates, along with the quantity and variety of educational resources available in the home, including electronic devices and books.

The teacher/educator's questionnaire consisted only of sections A, B and C. Caregivers of Italian monolinguals also completed the questionnaire, in which section C was aimed at exploring the family linguistic repertoire, to make it sure it was an Italian monolingual environment or whether regional languages were also spoken. The questionnaires were delivered during the Covid-19 pandemic, which made it complicated to reach all households and get all papers back. Therefore, the data obtained are incomplete:

Biographical and family data: dates and places of birth of all children are known. 54% (22/41) of families filled out the section in full, among them, 60% (18/30) of parents of bilingual children and 36% (4/11) of families of monolingual children provided complete information.

Schooling: information on the schooling of all children was obtained.

Language Exposure: 81% (33/41) of families filled out the section.

Home learning materials and activities: 61% (25/41) of families filled out the section.

3. Procedure

Children's narrative productions were collected between late 2021 (November and December) and early 2022 (January-March). Each bilingual participant took part in two storytelling sessions, which were videotaped.

Each session included a retelling task and a telling task. The first session was conducted in the family language with the help of four mediators: one for Spanish, one for Chinese, and two for Arabic (able to speak the various Arabic dialects), with the researcher being always present. The second session was conducted in Italian by the researcher with the assistance of a master student from the University of Bologna[60]. Monolingual children participated only in the Italian session.

To summarize, for each bilingual participant the following narratives were collected:

(i) retelling task in the family language

(ii) telling task in the family language

(iii) retelling task in Italian

(iv) telling task in Italian

[60] Despite being aware that randomization among participants constitutes the ideal scenario in administering the two tasks, due to constraints related to the availability of native speakers, it was not feasible to do so.

According to the instructions for the use of MAIN, the researcher should take care that the interval between the tasks in the family language and those in the societal language is at least 4-7 days, in order to minimize cross-linguistic transfer and memory effects. Efforts have always been made to comply with this indication during data collection. In order for the child to get familiar with the experimental setting and the adults, the opportunity for unstructured exploration of the room and engagement in enjoyable activities with the interlocutor was afforded, right before starting the task. A cozy and comfortable setting was prepared, consisting of a blanket to sit on, a teddy bear, a PC to play the story, and the printed pictures of the tale. This space was located within the school, and the activity took place during school hours.

Figure 3 Setting during the telling task.

As for the retelling task, the child had to look at the pictures of the narrative sequences (printed on paper) and follow the story, while a native speaker voice (previously recorded) was telling it (the audio was played using a PC). Then, the child was asked to retell the story, while looking at the pictures shared with the adult.

Three children did not produce any story in their family language (2 in Arabic and 1 in Chinese), possibly for emotional reasons. Originally, the plan was for the child to listen to story with an adult (no to leave the child alone) and then retell the story to the other adult who hadn't heard it. However, due to the stressful situation the child was experiencing, it was ultimately decided that both the child and the adult would engage in the listening process together and share pictures during the storytelling task.

In the telling task, the child observed the pictures from a different story, with respect to the story used in retelling, and was asked to narrate this story to the interlocutor, who sat in front of him/her with the images covered (Figure 3). Unlike the first session, there was no shared knowledge between the listener and the speaker, a factor that can influence the (in)definite status of the referent (Hickmann, 2003a).

Therefore, the analysis of NPs has been run only on the productions from the telling task, while all productions were used to construct linguistic profiles. Since the research focuses on the societal language, the analysis of NPs was conducted only on narratives in Italian; while those in the family languages have been used to get a complete view of the bilingual children competence, and to be able to establish dominance.

The working memory test was administered at the end of the first session and was presented as a game. The task included a naming component and a memory retention component. Each child was free to answer in the language he/she preferred, since the naming part did not contribute to the final score, which was calculated based on the number of items the child was able to retain in memory. Each session lasted maximum 30 minutes, due to COVID-19 restrictions imposed by the schools.

4. Data Analysis

Narratives were transcribed by a native speaker. Two analyses were performed on the Italian narratives: individual measures of language development were calculated from the narrative samples, both at the microstructural and macrostructural levels (Lindgren et al., 2023), and then all NPs (1248) were extracted and coded according to definiteness-related features. To assess linguistic dominance, the first type of analysis (individual measures of language development) was also performed on the family language texts. Each participant was assigned a score from the working memory test and child-external measures were calculated from the questionnaires. To summarize, for each participant the following data were collected:

(i) NPs in Italian narratives (for both bilinguals and monolinguals)

(ii) linguistic development measures in Italian (for both bilinguals and monolinguals)

(iii) linguistic development measures in the family language (only for bilinguals)

(iv) Working memory scores (for both bilinguals and monolinguals)

(v) child-external measures (for both bilinguals and monolinguals)

4.1 Narrative measures

Narrative analysis is a standard method for assessing vocabulary and grammar development (Castilla-Earls et al., 2021). For the analysis, all sentences were included except those with unintelligible parts. CLAN software (MacWhinney, 2000) was used for analysis. The following measures were obtained:

- **Sentence complexity:**
    - MLU5: the mean length in morphemes of each child's five longest utterances (MacWhinney, 2000), for both languages[61].
    - SubIndex: it is the ratio of the total number of main and subordinate clauses to the total number of sentences (Restrepo et alii, 2010). A sentence is defined as a main clause with all of its subordinate clauses and modifiers (Restrepo et al., 2010).[62]

- **Lexical diversity:**

  all the items produced in the language not under analysis were excluded.
    - VOCD (Vocabulary Diversity): it is derived from the TTR (Type/Token Ratio) versus token curve, which is computed using the transcript data, rather than relying on a

[61] In the literature, when selecting utterances with more than two morphemes, this kind of measure is also referred to as MMU (multimorphemic utterances) (see, for example, Genesee et al., 1996). In this study, we have reported the acronym and definition as used by MacWhinney (2000).

[62] Following Restrepo et alii (2010), all clauses coordinated by 'and', 'but', 'so', 'then' were coded as separate sentences. Subjectless clauses were also coded as new sentences (e.g., *e saltò*: "and [he] jumped").

specific TTR value within it. This measure offers three key advantages: it remains independent of the number of words in the sample, uses all available data, and provides more informative insights as it reflects how the TTR changes across a spectrum of token sizes. Studies have demonstrated that the VOCD measure outperforms previous measures in both avoiding the inherent flaw in raw TTR with varying sample sizes and in discriminating across a wide range of language learners and users (MacWhinney, 2000; Malvern et al., 2004)[63].

- NDV (Number of Different Verbs): since verbal lexicon is a very good predictor of grammatical development, the number of different verb types (roots) (NDV) was employed as a vocabulary measure (Hadley et al., 2016).

- **Dominance:**

In the assessment of linguistic dominance, participants' performance across three measures (MLU5, VOCD, NDV) is evaluated in both Italian and their family language. SubIndex was excluded, because the strategies of coordination and subordination are very different interlinguistically and therefore difficult to compare. It follows that this measure should only be used in comparing bilingual children to monolinguals. Each measure is examined individually to determine whether the participant achieved a higher score in Italian or their family language. Subsequently, the frequency of higher scores in Italian and the family language is tallied. Finally, the language with the highest number of measures where it achieved the superior score is deemed dominant (see Kupisch, 2007).

[63] See McCarthy and Jarvis (2007) for a discussion on this lexical diversity measure.

4.2 Working Memory

The "Tieni a Mente" test of the FE-PS 4-6 battery (Usai et al., 2017) was administered for working memory (WM) assessment to all participants. The test includes a naming task and a retaining task; the answers given in the naming task do not contribute to the scoring. 1 point is assigned for each correct answer. Half a point is given for the correct answer that is provided only after the box with the figures of the required category is shown (the figure supports the child in retrieving the correct item). In the case of an incorrect answer or missing answer, zero is attributed. The total score is calculated by summing the scores obtained in six series (see 4.3 above), for a maximum of 9 points. High scores indicate a greater ability to update information within the working memory.

4.3 Child-external measures

Based on the literature (Paradis, 2011; Unsworth, 2016 among others), four major external factors have been considered in the analysis: SES, quantity of input, quality of input, home literacy activities.

1) SES (socio-economic status): the socioeconomic status of the families was calculated based on the occupation and educational level of both parents. Levels of education were derived from the International Standard Classification of Education ISCED (UNESCO, 2011), and the scores assigned are presented below:

   Level of Education (ISCED-11)[64]: none (1 point), Primary School (2 points), Secondary School (3 points), University (4 points).

[64] For the sake of economy of analysis and to keep the questionnaire easily accessible, the ISCED-11 number of levels of education was reduced from 9 to 4: "early childhood education" was merged with the "none" level; the label "secondary school" was maintained, but the distinctions between "lower" "upper" and "post-secondary non-tertiary education" was eliminated; post-secondary education (from short-cycle tertiary education to PhD) was collapsed under the more intelligible label of "university."

The International Standard Classification of Occupations ISCO (ILO, 2008) was followed to classify parents' jobs. In this classification, occupations are organized into four groups according to so-called skill levels, "defined according to the complexity and range of tasks and duties to be performed in an occupation" (ILO, 2008[65]). The scores given are presented below:

Level of Occupation (ISCO-08): skill 4 (5 points), skill 3 (4 points), skill 2 (3 points), skill 1 (2 points), not occupied (1 point).

2) QUANTITY OF INPUT IN ITALIAN: two quantity-oriented measures have been computed, based on the number of sources providing input in Italian and on the frequency of activities conducted with native Italian language input providers:

Input sources: for each source of input, included in the questionnaires (mother, father, siblings, friends from the same linguistic background, monolingual Italian-speaking friends, cousins, and other relatives), a scoring system was employed. In this system, for each source 1 point was assigned if the child exclusively received input in the family language from that source, 2 points if the source provides input in both languages, and 3 points if input is exclusively received in Italian. 0 corresponds to absence of the source.

Frequency of activities in Italian: for each activity that was carried out in Italian a score of frequency was assigned from 4 (every day), 3 (once a week), 2 (less than once a week), to 1 (never).

[65] https://ilostat.ilo.org/resources/concepts-and-definitions/classification-occupation/
https://www.ilo.org/public/english/bureau/stat/isco/isco08/index.htm

3) QUALITY OF INPUT IN ITALIAN: two quality-oriented measures were calculated. These measures consider Italian language learning activities in which the child is involved and the self-reported proficiency level of the parents in Italian.

Activities in Italian: caregivers were asked which language the child typically uses while engaging in certain activities (watching TV or other media, playing with friends at home, playing with friends outside, spending time with parents outside). A score of 1 was assigned to activities carried out exclusively in the family language, 2 points for activities conducted in both languages, and a score of 3 points for activities performed solely in Italian. 0 corresponds to absence of engagement in the activity.

Parents' Proficiency in Italian: parents were required to assess their Italian language skills in speaking, listening, reading, and writing on a scale ranging from 4 (not at all) to 1 point (very good).

4) HOME LITERACY: this measure computes the frequency of activities that may fall under the label of home literacy, (Barton & Hamilton, 2000; Cairney, 2009; Torregrossa & Carbonara, 2022) such as e.g., counting, reading aloud, pretending to write, learning ditties and nursery rhymes, actively listening to, or telling stories, regardless of the language used. For each activity the respondent was asked to indicate the frequency: "Always" (4 points), "Often" (3 points), "Sometimes" (2 points) to "Never" (1 point). In addition to these, yes (1 point) - no (0 points) questions are asked, about the presence of electronic devices (PCs, smartphones, tablets) and the number of books at home (from less than 10 = 1 point to two shelves or more/more than 100 books = 4 points) (see INVALSI, 2009 for a similar scale).

4.4 Transcription

Each narrative was transcribed. The stories in Italian and Spanish were transcribed by the researcher, while for Arabic dialects and Chinese native speakers were selected among university students and PhD candidates. They were trained on transcription conventions. The transcription guidelines, as provided by MAIN, were followed (MAIN team, 2021). Transcriptions are not publicly available at the moment. Each participant has been anonymized, and individuals will be referred to through the use of acronyms composed as follows: CHILD + a progressive number from 1 to 11 based on the order of interview in Italian + acronym for the language spoken at home: IT = Italian monolingual speakers, ARA = Italian-Arabic bilinguals, SPA = Italian-Spanish bilinguals, CHI = Italian-Chinese bilinguals (e.g, CHILD3ITA, CHILD3ARA, CHILD3SPA, CHILD3CHI).

4.5 NP coding

All NPs were extracted from the narratives in Italian (for a total of 1248 NPs). Those produced in the family language and those not related to the narrative (comments, interactions with the interlocutor on topics other than the story) were excluded. Each NP was coded according to definiteness-related features:

- Referentiality: this feature assumes two values (see 1.1.1), the NP can be:
    - Referential
    - Non referential
- Identifiability: if the NP is referential, then it can be:
    - Identifiable: the speaker assumes the hearer can pick out, from all the referents that might be categorized in this way, the referent he/she has in mind (Chafe, 1976)
    - Non-identifiable: the speaker cannot make this assumption. In the corpus, cases of first mention of a referent are the only instantiations of this type of referents.

- If the referent is identifiable, 4 types of sources of identifiability are possible (see 1.3.1):
    - Anaphor: the referent is identifiable because it is already introduced in the previous discourse
    - Associative: the referent is identifiable because it is associated to a referent already introduced in the previous discourse (association can be kinship, possession or frame-induced)
    - General knowledge: the referent is identifiable because it is part of the encyclopedic knowledge shared by the interlocutors
    - Situational: the extralinguistic context is the source of identifiability.

This level of analysis indirectly involves animacy as well, as referents that are part of general knowledge, inalienable possessions, and those that are frame-induced are often inanimate and frequently mentioned for the first time as definite and belong to the background of the scene (see examples below). Anaphoric NPs, on the other hand, often refer to characters (thus animated) that should be introduced as new (non-identifiable) and then referred to with the definite article in subsequent mentions[66].

- Type of determiner: this feature takes on four values:
    - bare nominal
    - definite article
    - indefinite article

[66] It must be acknowledged that other linguistic categories interact with definiteness, such as, individuation, agency, and discourse prominence or salience, which have not been directly included in the coding system to avoid making the analysis overly complex and to limit it to the core function of definite expressions: identifiability. For a discussion on the interaction between definiteness and various semantic categories, see Khan (1984), Croft (1990:116), Lyons (1999: 213-215), among others.

- demonstrative[67]: it is the only determiner considered here, besides the article. Taking demonstratives into consideration is useful for the study of cross-linguistic influence, since in some languages they overlap with certain functions of articles, while in others, they may co-occur with them (see 1.4.4).

- Target-like: a form is defined as a target if it is adult-like in terms of
  - Grammaticality of the use of bare forms instead of determiner + noun
  - Appropriateness of form to the identifiability status of the referent

Beyond the categories driven by pragmatic considerations (appropriateness), the only syntactically motivated category examined in this study for all groups was the presence/absence of the determiner (grammaticality), which proves crucial in evaluating the impact of the family language, especially with articleless languages like Chinese.

As Chinese heavily relies on constituent order to convey referent identifiability status, to test the hypothesis of cross-linguistic influence from the family language, an additional level of syntactic coding was added only for the productions of ITA-CHI bilingual children. Each NP was tagged for syntactic role (subject, object, and oblique), and those serving as subjects or objects were labeled based on their position relative to the verb: preverbal or postverbal.

4.5.1 Examples of analysis

Some examples from the corpus are given below to make the coding system clearer.

First, target examples of determiner types in appropriate contexts will be provided (under each example English word-by-word translation is given in italics[68]). The first example provides an NP with the indefinite article (1):

[67] In the coding system, distal and proximal demonstratives were not distinguished, as that distinction falls outside the scope of this study.
[68] In the translation, morphological errors are not highlighted, e.g., lack of agreement between determiner and noun in gender and number, errors in verbal morphology, etc., as they are not the primary focus of this work.

1) CHILD1IT: c'era **un uccello** nero.
*there was **a black bird.***

In this case the NP is referential, the referent is non-identifiable since it is its first mention. It is introduced with an indefinite article, which is the target form.

The cases with the target definite article follow, as they are in an identifiable context (2-6):

2) CHILD1IT: mentre **l'uccello nero** ha morso la coda alla volpe. (bird already mentioned)
*while **the black bird** bit the fox's tail.*

In this example the NP is referential, the referent is identifiable, since it has already been mentioned in the previous discourse, it is a case of anaphoric usage. It is introduced with a definite article, which is the target form.

3) CHILD10IT: con **la testa** lo stava riprendendo. (referring to the father who saves his son)
*with **the head** he was taking him back.*

In this case the NP is referential, the referent (*the head*) is identifiable, as it is associated with another referent (*the father*) already mentioned in the previous discourse: this is a case of identifiability by association. The referent is introduced with a definite article, which is the target form.

4) CHILD2IT: [picture 1] c'è una pecorella che è annegata …
[picture 4] dopo le altre pecorella sono venute a bere **dal fiume**.
*there is a sheep that drowned ... after the other sheep came to drink from **the river.***

In this example the NP is referential, the referent (*the river*) is identifiable, as an element belonging to a frame that has already been activated in the previous discourse by the verb *drown*, which is the anchor (for frame-induced identifiability see 1.3.1). This case is also placed under the label "associative". The referent is introduced with a definite article, which is the target form.

5) CHILD5ARA: una pecorella è caduta sul fiume e **la mamma** si è spaventata.
*a little sheep fell on the river and **the mother** was frightened.*

In this utterance the NP is referential, the referent (*the mother*) is identifiable, as it is associated with another referent (*a little sheep*) already mentioned in the previous discourse: this is also a case of identifiability by association. The referent is introduced with a definite article, which is the target form.

6) CHILD5SPA: e l'altra pecora stava mangiando **l'erba.**
*and the other sheep was eating **the grass**.*

In (6) the NP is referential, the referent (*the grass*) is identifiable, as it is part of the encyclopedic knowledge: this is a case of identifiability based on general knowledge. In the narratives examined, these cases are limited to mass nouns such as 'grass' and 'water'. The referent is introduced with a definite article, which is the target form.

Given the experimental setting, where the pictures were not available to the interlocutor, uses in which the extralinguistic context (i.e., pictures) makes the referent identifiable were not expected. Nevertheless, "situational uses" were included in the coding system in order to identify all those NPs produced in conjunction with explicit calls for the interlocutor's attention to the picture (expressions like "look!" or pointing gestures). In the absence of clear gestures or expressions that draw the interlocutor's attention to the picture, it is impossible to determine whether determiners in the other contexts are also triggered by the images.

Now an example with the demonstrative is provided (7):

7) CHILD3ITA: poi allora una volpe molto affamata vide **quelle pecorelle.**
*then a very hungry fox saw **those sheep.***

In this case the NP is referential, and the referent is identifiable, as it has already been mentioned in the previous speech, the source of identifiability is anaphora. It is introduced with a demonstrative, which is target.

We now move on to examples of bare nominals used grammatically (8 a, b, c):

8) a. CHILD1ITA: poi le due caprette sono **in acqua.**
*then the two little goats are **in (the) water.***
b. CHILD4CHI: e un mucca sta bevendo **acqua.**
*and a cow is drinking **water**.*
c.CHILD5CHI: e mucca piccola mangia **foglie**
*and cow little eats **leaves.***

The utterances in (8) are examples of target bare nominals, that is, cases in which the grammar of Italian allows the noun without a determiner. Specifically, in (8a) there is a mass noun in a PP, in an identifiable context by general knowledge; in (8b) the mass noun is object, and its referent is also identifiable by general knowledge; example (8c) shows a case of a bare plural countable noun, with a non-identifiable referent (see 2.1).

We now proceed to show examples of NPs annotated as non-target. We will start with the indefinites (9 a, b), then we move on to present examples with the definites (10), with the demonstratives (11), and with the non-target bare nominals (12 a, b). Finally, examples of non-referential NPs where the opposition of identifiability does not apply will be given (13):

9) a. CHILD1ITA: **una capretta** abbracciava tutte le caprette.
***a little goat** hugged all the little goats.*

(uttered at the end of the story, referring to the main character already mentioned)

b. CHILD3SPA: e poi una **una mamma** ha aperto la bocca.
*and then **a mom** opened the mouth.*

The cases in (9) are examples of nontarget indefinite article. In (9a) the NP is referential, and the referent is supposed to be identifiable, since it has already been mentioned in the previous discourse several times (it is a case of anaphoric usage), therefore the indefinite article is not the expected form. The NP in (9b) is also referential, and the referent (*mom*) is identifiable; in fact, as in (5) this is a case of identifiability by association. The use of the indefinite with a uniquely identifiable referent in the

context of utterance is misleading and might lead the hearer to think that there is more than one mom in the story, which is not the case.

10) a. CHILD10CHI: c'è **l'uccellino.** (the character is introduced for the first time)
*there was **the little bird**.*

In this case the NP is referential; the referent is not identifiable since it is mentioned for the first time. However, the child introduces it with a definite article, which is not the target form.

11) CHILD4ARA: poi **quel piccione** vede la volpe. (the bird is introduced for the first time)
*then **that pigeon** sees the fox.*

In this example the NP is referential. The referent is not identifiable, as it is mentioned for the first time. However, the child introduces it with a demonstrative, which is not the target form.

12) a. CHILD5CHI: poi **uccellino** vede.
*then **bird** sees.*
b. CHILD4CHI: e dopo c'è una cagnolino che guarda **mucca.**
*and after there is a little dog that looks **cow.***

The examples in (12) show ungrammatical bare nominals. In these cases, the NP is not target, because, regardless of the identifiability of the referent, the grammar of Italian requires a determiner. Since the bird and the cow were already introduced in the discourse, they were coded as NPs in identifiable context (anaphoric usage).

13) a. CHILD9CHI: è **una mucca.**
*(it) is **a cow.***
b. CHILD1SPA: e la mamma capra e piccolo capre le vuole dare **un bacio.**
*and the mother goat and little goat want to give her **a kiss**.*

The examples in (13) show nonreferential NPs. In the first statement there is a predicative (or qualitative, see 1.1.1) use of the NP while in the second statement it is a nonspecific NP (see 1.1.1).

4.5.2 Special cases

In the previous cases, the coding procedure was straightforward and easy to evaluate through objective criteria, given the simplicity of the distinction between identifiable and non-identifiable contexts in children's storytelling. However, the cases we will discuss below posed more challenges. This difficulty arises from the issue mentioned in 1.3, that is, the definite-indefinite contrast is binary, it is a matter of yes or no, whereas identifiability is a matter of continuum, in which different degrees can be identified.

The solutions adopted in those cases will be exposed below:

14) a. CHILD8ITA: c'è una caprettina **nel fiume** (first mention of the river)
*there is a little goat in **the river**.*
b. CHILD9ITA: una volpe dietro **l'albero** li vedeva. (first mention of the tree)
*a fox behind **the tree** saw them.*
c. CHILD6ARA: quando ha visto era nascosto vicino al**l'albero.** (first mention of the tree)
*then when he saw it was hidden near **the tree.***

All the examples in (14) present instances of spatial elements being introduced for the first time, during the unfolding of the story, with a definite article. To be introduced for the first time as definite (i.e., identifiable to the hearer) this kind of items should be anchored in a frame that must have already been activated for the hearer and be well-defined (Du Bois, 1980; Schwarz, 2013; see (4), instead, for a case of a clearly activated frame). Here, we chose to code as non-identifiable those contexts in which it was not possible to find an unambiguously already activated frame, which the referent had to anchor to. This is the case of all the utterances in (14), which are of therefore nontargets.

15) a. CHILD1ITA: la volpe tirava per **una zampa** la capretta.
*the fox pulled by **a paw** the baby goat.*
b. CHILD8ITA: poi la volpe ha preso **la zampa** dell'agnello.
*then the fox took **the hoof** of the lamb.*

The utterances in (15) refer to the same episode, described by two different children. These are examples of a partially identifiable referent, that stands in a one-to-many relationship with the anchor referent (*baby goat* and *lamb*: one animal-four hooves). When presented with the picture, the child can encode the referent 'paw' in two distinct ways: as a non-identifiable entity within a set of four items of the same type, realized with a specific indefinite article, or as identifiable via bridging (we all know that lambs/goats have hooves). What is identifiable in the second case is the type of referent. In both cases it is unnecessary to specify which of the four hooves is being taken (recall Du Bois, 1980's 'curiosity principle', seen in 1.3.1). Both cases were considered as target realizations.

16) a. CHILD6SPA: che voleva mangiare quella pecora che stava mangiando **dell'erba.**
*who wanted to eat that sheep that was eating **some grass**.*
b. CHILD7ARA: hanno visto un'altra pecora che stava mangiando **dell'erba.**
*they saw another sheep that was eating **some grass**.*

Those reported in (16) are the only two occurrences of partitives used with a mass noun. In these two cases the context was annotated as non-identifiable and the form as indefinite and target. In this way the indefinite form was mapped on non-identifiable context and is therefore target (for a discussion on the determiners with mass nouns and their respective interpretations in Italian, refer to section 2.1).

17) a. CHILDITA6: ci sono anche **dei fiori.**
*there are also **some flowers**.*
b. CHILSPA7: c'erano una volta **delle pecorelle.**
*once upon a time there were **some sheep**.*

Cases of partitives, like those in (17), are few in the corpus and, as we have already illustrated in 2.1, are to be considered as plural forms of the indefinite article. In the two cases presented above, NPs with the partitive occur in non-identifiable contexts, for which they were considered target forms.

No attention was paid to the gender and number agreement between the determiner and the NP, as these aspects fall outside the interest of this study. Consider the following example:

18) CHILD6CHI: **la lupo** andato là.
the.F.SG wolf.M.SG *gone there.*

In this case the definite article is in the feminine singular form, while the noun is masculine singular: the agreement is correct only for the number category. However, these morphological aspects are disregarded in the evaluation of the data: since this was a NP in identifiable context via anaphora, the use of definite article is considered target.

NPs with determiners other than articles and demonstratives were not considered in the present analysis. Specifically, the following phrases were excluded: numeral + noun, possessive + family name (see 2.1). Possessives and numerals are part of complex definites (see 1.4), meaning they interact with the definiteness category, but beyond this, they encode other notions such as quantification, possession, etc., and have more complex syntactic realizations (see 2.1, examples 43a and b). This would have introduced additional complexity both in the coding system and analysis, going beyond the scope of the current work, which focuses on encoding identifiability.

5. Results and Discussion

In the preceding chapters, I discussed the theoretical foundations and the methodologies of the research project. This chapter aims to provide an account of the empirical findings gathered through data collection and analysis. These findings address the central research questions presented in the previous chapter, reiterated below:

1. Do bilinguals and monolinguals distribute articles differently according to identifiability?

2. Does family language influence the societal language in the acquisition of definiteness?
3. Is there an effect of child-external and internal factors on the acquisition of definiteness?

To address these inquiries, narrative productions in Italian were examined (using the "Baby goats" story). The narratives were elicited through the telling task, which was administered to all four groups of children. The analysis is primarily quantitative in nature (conducted using R software; R Core Team, 2021). This chapter is organized in four sections. Section 1 describes the distribution of determiners according to identifiability in the four groups of children. Section 2 is about the effect of the family language on Italian, looking at divergent patterns. Section 3 is dedicated to the impact of child-internal and child-external factors on the expression of definiteness. In Section 4 summary and discussion are provided.

1. Encoding identifiability

The first question to be addressed in this section is: do bilinguals and monolinguals distribute articles differently according to identifiability?

Below, the statistics to describe the distribution of forms (definite articles, demontrstives, indefinite articles, bare nouns) in the four linguistic groups are presented: Italian monolinguals (hereinafter ITA), Italian-Arabic speaking bilinguals (ITA-ARA henceforth), Italian-Chinese speaking bilinguals (hereinafter ITA-CHI), Italian-Spanish speaking bilinguals (ITA-SPA henceforth).

In order to answer the question, it is necessary to know whether there are differences by group according to context. Two contexts have been identified based on the identifiability status of the referent: the identifiable context (Id henceforth) and the non-identifiable context (NId henceforth) (see Chapter 1 for definitions). Table 1 shows the number of NPs produced in the two different contexts, in the four language groups. The non referential noun phrases (NR), on which the distinction between identifiable and non-identifiable is not applicable, are also included.

A quick look at the table clearly shows that in these narratives most NPs appear in identifiable contexts, which in all groups is around 80%, compared with around 20% of nouns produced in non-identifiable contexts. It is clear from the table that there are no major differences between the groups in any of the contexts. A chi-square analysis revealed that the association between language and context was not significant (χ^2 = 1.5177, df = 3, p = 0.67). This shows that the narratives produced by children from different language groups are indeed comparable.

Context	ITA	ITA-ARA	ITA-CHI	ITA-SPA	Total
ID	256 (79%)	243 (82%)	233 (78%)	264 (83%)	996 (81%)
NID	65 (20%)	51 (17%)	52 (17%)	53 (17%)	222 (18%)
NR	1 (0,3%)	2 (0,7%)	13 (4%)	1 (0,3%)	17 (1%)
Total	322 (100%)	296 (100%)	298 (100%)	319 (100%)	1235 (100%)

Table 1 Frequency of NPs (with % in parenthesis), by context and language groups.

Given that there are no significant differences in the story children told and the distribution of contexts among the four groups, the next question I addressed is: do the groups know the main formal strategies required in identifiable and non-identifiable contexts in Italian? Since the non referential contexts are very few, they have been excluded from the following analysis.

I now turn to the distribution of NP types in identifiable contexts by each group, which is shown in Table 2.

DETERMINER	ITA	ITA-ARA	ITA-CHI	ITA-SPA	Total
Definite art.	239 (93%)	219 (90%)	51 (22%)	248 (94%)	757 (76%)
Demonstrative	7 (3%)	10 (4%)	25 (11%)	1 (0,4%)	43 (4%)
Indefinite art.	5 (2%)	5 (2%)	23 (10%)	2 (1%)	35 (4%)
Bare nouns	5 (2%)	9 (4%)	134 (58%)	12 (5%)	161(16%)
Total	256 (100%)	243 (100%)	233 (100%)	264 (100%)	996 (100%)

TABLE 2 Frequency of determiner (with % in parenthesis) in identifiable contexts by language groups.

Both for monolingual Italians, ITA-ARA and ITA-SPA bilinguals, the most used determiner in identifiable contexts is the definite article: children show respectively 93%, 90% and 94% of frequency, while ITA-CHI group shows a much lower frequency (22%). Below, examples (19-22) contain definite articles used in identifiable contexts by all four groups (English word by word translation is provided for each example):

19) CHILD1ITA	20) CHILD1ARA	21) CHILD10CHI	22) CHILD7SPA:
il corvo inseguiva <u>la volpe</u>	**l'uccellino** voleva prendere <u>il cane</u>	**l'uccellino** prendere <u>la cane</u>	**il corvo** stava inseguendo la volpe
the crow *chased <u>the fox</u>*	***the little bird*** *wanted to catch <u>the dog</u>*	***the little bird*** *to catch <u>the dog</u>*	***the crow*** *was chasing <u>the fox</u>*

The NPs in bold are extracted from descriptions of the same image. In the story children were telling, both the characters, namely the crow and the fox, had already been introduced in the previous episode and coded with a definite article, resulting in a target-like choice.

Only the group of Chinese-speaking bilinguals shows a strong preference for bare nouns in identifiable contexts, with a frequency of 58% (in ITA-CHI group around 86,6% of determinerless NPs occurs in contexts where determiners are obligatory, resulting in ungrammatical NPs). Bare nominals are very rare in the narratives of Italian-speaking monolingual children (2%, all instances are grammatical), the frequency is also very low in Spanish-speaking bilingual children (5%, 8 out of

13 instances are ungrammatical) and their frequency is slightly lower among Arabic speakers (4%, 4 out of 9 instances are ungrammatical).

Examples (23-27) show grammatical instances of bare nouns from all the groups (23-26 are about the same picture):

23) CHILD11ITA	24) CHILD7ARA	25) CHILD3CHI	26) CHILD7SPA:
e un caprettino che era **in acqua**	è scesa **in acqua**	il suo bambino pecora è **in acqua**	il piccolo capra è andato **in mare**
*and a little goat that was **in (the) water***	*(she) went down **in (the) water***	*her baby sheep is **in (the) water***	*the little goat went **into (the) sea***

27) CHILD5CHI
questa mucca mangia **erba**
*this cow eats **grass***

Note that the overwhelming majority of grammatical bare nouns in the narratives analyzed are mass nouns in PP (as in 23-26). Instances of grammatical bare mass nouns in argument position (as in 27) are produced mostly by Chinese-speaking children.

Examples (28-30) show ungrammatical bare nouns from the three bilingual groups:

28) CHILD10ARA	29) CHILD10CHI	30) CHILD3SPA:
poi **uccello** è arrabbiato tanto	e **pecora** vuole scappare	**lupo** ha preso piedi della pecora
*then **bird** is very angry*	*and **sheep** wants to run away*	***wolf** took feet of the sheep*

Recall that bare countable nouns in subject position, as in (28-30 in bold) or in object position as in (30, underlined), are not allowed in Italian (see chapter 2.1). The majority of occurrences of ungrammatical bare nouns can be attributed to only one individual per group: CHILD10ARA and CHILD3SPA. The former is a 7-year-old girl classifiable as an early L2 learner, as her exposure to Italian began at the age of 6, while the latter is the youngest participant in the Spanish-speaking group (4;1 y.o.) and is a simultaneous bilingual.

Use of demonstratives is generally low in three groups: ITA 3%, ITA-ARA 4%, ITA-SPA 0,4 %, while it is higher, around 11%, in the ITA-CHI group. Examples are given below:

31) CHILD3ITA	32) CHILD7ARA	33) CHILD3CHI	34) CHILD1SPA:
poi allora una volpe molto affamata vise **quelle pecorelle**	ma voleva mangiare tutte **quelle pecore**	e **questa pecora** sta mangiando erba	la volpe vuole mangiare **la capra questa**
*then a very hungry fox saw **those little sheep***	*but he wanted to eat all **those sheep***	*and **this sheep** is eating grass*	*the fox wants to eat **the goat this***

Examples (31-32) are considered contextually adequate, since they are instances of anaphoric NPs referring back to elements already present in the previous discourse. For (33) and (34) see section 2, since they can be seen as results of cross-linguistic influence.

In identifiable contexts, the number of indefinites is negligible (ITA 2%, ITA-ARA 2%, ITA-SPA 1%). Also in this case, the group of Chinese-speaking bilinguals is distinguished from the others by a higher frequency, around 10%. Examples (35-38) are from episodes where all the referents had already been introduced and were supposed to be identifiable to the interlocutor, nevertheless an indefinite article is used:

35) CHILD1ITA	36) CHILD3ARA	37) CHILD7CHI	38) CHILD2SPA:
una capretta abbracciava tutte le caprette	e dopo il lupo l'ha preso la coda **uno uccellino**	**un volpe** saltò per mangiare un capretto	è venuta la volpe che ha visto **una capretta**
***a little goat** hugged all the goats*	*and after the wolf took his tail **a little bird***	***a fox** jumped to eat a kid*	*the fox came and **saw a little goat***

Those few instances of indefinites in identifiable contexts (the so-called discourse-integration errors) are all concentrated in anaphoric contexts. No errors occur in the other sub-types of identifiable contexts (bridging, shared knowledge, situational; see Appendix C for a table with counts and percentage of sub-types of identifiable contexts).

A chi-square analysis was then conducted in order to compare each bilingual group to the control group of monolinguals. It revealed that the differences in the selection of type of determiner

in identifiable contexts are significant for the Chinese speaking group ($\chi^2 = 262.79, df = 3, p < .001$) and for the Spanish speaking group ($\chi^2 = 9.3867, df = 3, p = 0.02$) , but not significant for the Arabic speaking group ($\chi^2 = 2.2085, df = 3, p = 0.53$).

The distribution just described shows that monolingual children know the definite article is the main formal strategy for expressing identifiability in Italian, followed by demonstratives, with very few cases of substitution (discourse-integration errors: indefinite article instead of definite article). The same can be said of the bilingual groups of Spanish and Arabic speakers. The Chinese-speaking bilingual children behave differently, as they often omits the definite article in identifiable contexts, producing a large number of bare nominals. Also demonstratives and cases of substitution are more frequent in this group than in the others.

The analysis of the residuals indicates that the asymmetry between monolinguals and Chinese-speaking children is concentrated on bare nouns (see mosaic-plots in Appendix C). No major differences in the use of determiners in identifiable contexts between monolinguals and the other two bilingual groups emerge from the analysis of residuals (see mosaic-plots in Appendix C).

We now turn to the distribution of NP types in non-identifiable contexts by each group, which is shown in Table 3.

DETERMINER	ITA	ITA-ARA	ITA-CHI	ITA-SPA	Total
Definite art.	17 (26%)	17 (33%)	9 (17%)	18 (33%)	61 (27%)
Demonstrative	0 (0%)	2 (4%)	1 (2%)	1 (4%)	5 (2%)
Indefinite art.	48 (74%)	30 (59%)	28 (54%)	34 (63%)	140 (63%)
Bare nouns	0 (0%)	2 (4%)	14 (27%)	0 (0%)	16 (7%)
Total	65 (100%)	51 (100%)	52 (100%)	54 (100%)	222 (100%)

Table 3 Frequency of determiner (with % in parenthesis) in non-identifiable context by language

In non-identifiable contexts, the most used determiner is the indefinite article in all groups, albeit with different frequencies between monolinguals and bilinguals: 74% among monolinguals, 59% in Arabic-speaking, 54% in Chinese-speaking and 63% in the Spanish-speaking bilingual children. See

examples (39-42) below, all used to introduce a new referent in the story, namely the black bird, resulting in adequate referential uses:

39) CHILD11ITA	40) CHILD9ARA	41) CHILD7CHI	42) CHILD2SPA:
arrivò **un corvo**	dopo è arrivato **un uccellino**	c'è **un uccello** sul ramo	poi viene **un corvo**
a crow *came*	*then* ***a little bird*** *arrived*	*there is* ***a bird*** *on the branch*	*then comes* ***a crow***

The frequency of definites in non-identifiable contexts is 26% in monolinguals, 17% in Chinese-speaking, 33% in both Spanish and Arabic-speaking bilingual children. As the previous examples (39-42), those in (43-46) were all used to introduce new referents in the story, namely the black bird and the fox, but utilizing definite articles, resulting in egocentric errors (definite article instead of indefinite article):

43) CHILD10ITA	44) CHILD5ARA	45) CHILD10CHI	46) CHILD8SPA:
e nel frattempo **il corvo** lo stava guardando	e c'era **la volpe** affamata di pecore	c'è **l'uccellino**	**la volpe** che si è nascosta vise le caprette
and meanwhile ***the crow*** *was watching it*	*and there was* ***the fox*** *hungry for sheep*	*there is* ***the bird***	***the fox,*** *which hid, saw the goats*

There is also a peculiar class of referents in the narratives that is worth mentioning here, since it is mostly encoded with a definite article when mentioned for the first time. These are spatial referents, functioning as the main spatial anchor points, to locate the characters (note that ITA-CHI group prefers bare nouns for this kind of referents). Children typically introduce these elements, which initially function as background components within the setting, when they become focal points to which the actions of the central character in the episode are anchored. Their identification is often left to the presupposition of the listener, not always resulting in discourse-pragmatic adequate choices. See the following examples (47-50) which show the first mention of two of the spatial anchors present in the stories, namely the tree and the lake, coded with a definite:

47) CHILD9ITA	48) CHILD2ARA	49) CHILD8CHI	50) CHILD8SPA:
una volpe dietro **l'albero** li vedeva	poi è venuta una volpe dietro al**l'albero**	una mucca mangia **l'albero**	è andato proprio giù **sul lago**
*a fox behind **the tree** saw them*	*then a fox came behind **the tree***	*a cow eats **the tree***	*he went right down to **the lake***

In non-identifiable contexts, monolingual Italian speakers do not use demonstratives, but bilinguals occasionally do: 4% in bilingual Spanish and Arabic speakers, and 2% in Chinese speakers. See examples (51-53):

51) CHILD4ARA	52) CHILD8CHI	53) CHILD6SPA:
poi **quel piccione** vede la volpe	un mucca sta mangiando **quella albero**	poi c'era un cane che voleva mangiare **quella pecora** che stava mangiando dell'erba
*then **that pigeon** sees the fox*	*a cow is eating **that tree***	*then there was a dog that wanted to eat **that sheep** that was eating grass*

In these examples the referents have never been introduced before, nor could be identified by the interlocutor, who did not see the images, via other sources. Note that the bird in (51) is mentioned in the story for the first time, the example (52) is the beginning of the story (therefore nothing suggests the presence of a tree to the listener), and in (53) no eating animals were mentioned before, resulting in NPs for which the child assumes a greater common ground than the actual one.

Bare nominals in non-identifiable contexts occur among Chinese-speaking bilinguals, with a frequency of 27% and among Arabic-speakers, with a frequency of 4% (see examples 54, 55 and 56 below). They don't occur at all neither among monolinguals nor among Spanish speakers. It is worth recalling that in the coding system adopted in this study, only bare plurals were considered adequate in non-identifiable contexts. Mass nouns can also be bare in Italian, but in the coding system they were considered identifiable via shared knowledge (refer to chapter 2 for theory and chapter 4 for the coding system). See examples (54-56) showing bare nouns from two bilingual groups (ITA-ARA and ITA-CHI):

54) CHILD9ARA	55) CHILD10CHI:	56) CHILD5CHI
c'era **pecora** che è andata nell'acqua	la pecora sta mangiando **foglie**	poi **uccellino** vede
*there was **sheep** that went into the water*	*the sheep is eating **leaves***	*then **little bird** sees*

(55) is an example of a grammatical and contextually adequate bare plural in a non-identifiable context. Examples of grammatical bare plurals are produced only by ITA-CHI speakers. Examples 54 and 56 show ungrammatical determinerless NPs lacking the indefinite article (the referents were introduced for the first time). In this case as well, most of the examples of ungrammatical bare nouns can be attributed to CHILD10ARA.

According to the distribution showed in Table 3, both monolingual and bilingual children know that in Italian in non-identifiable contexts, the preferred determiner is the indefinite article. Cases of substitution, that is, use of the definite article with non-identifiable referents (egocentric errors), are more frequent than cases of indefinites in identifiable contexts (discourse-integration error). In the Chinese-speaking group, article omission is more frequent than article substitution. The frequency of bare nominals in non-identifiable contexts in the Chinese-speaking group is close to the frequency of the use of definite article in the monolingual group.

A chi-square analysis was conducted in order to compare each bilingual group to the control group of monolinguals. It revealed that the association between language and type of determiner in non-identifiable contexts is significant for the Chinese speaking group ($\chi^2 = 21.546$, $df = 3$, $p < 0.001$), but not significant neither for the Arabic speaking group ($\chi^2 = 6.5597$, $df = 3$, $p = 0.08$), nor for the Spanish speaking group ($\chi^2 = 2.2215$, $df = 2$, $p = 0.32$).

Like in the previous case, the residual analysis shows that the most pronounced asymmetry with monolinguals is in the use of bare nouns by the CHI-ITA group (see mosaic-plots in Appendix C). Differences in the use of other determiners in the other two groups are not significant (see mosaic-plots in Appendix C). However, an asymmetry can still be noted distinguishing the control group from the experimental group of bilingual children in the use of indefinites. In order to visualize this

contrast between monolinguals and bilinguals, in Figure 4 the distribution of indefinites in the two contexts for each language group is represented:

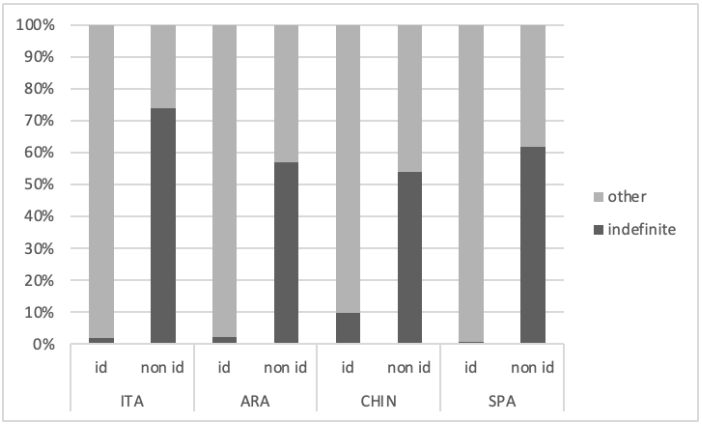


Figure 4 Distribution of indefinites, by context and by language group.

The figure shows a clear preference for the use of indefinites in non-identifiable rather than identifiable contexts in all groups. However, this preference is more accentuated in monolinguals than in the three bilingual groups. This could suggest slightly greater difficulty in encoding the non-identifiable status of the referent by bilingual children, compared to monolingual peers. It should be underlined, however, that although visibly clear from the figure, the difference with the monolinguals in the coding of non-identifiability reached the level of significance only for the ITA-CHI group.

To sum up, the monolingual group shows clear context sensitivity and knowledge of the formal strategies to be used in the appropriate context. The groups of bilinguals are also sensitive to context, but with different distribution of forms depending on the family language. The ITA-ARA and ITA-SPA groups seem very similar to each other and also behave similarly to their monolingual peers, with instances of ungrammatical omissions that monolinguals do not show. ITA-CHI speaking group appears to be different from all the other groups. Chinese-speaking bilingual children seem to be able to distinguish identifiable referents from non-identifiable ones, preferring bare nouns or demonstratives for the former and indefinite articles (with instances of omission) for the latter.

2. Effect of family language in divergent patterns

In this section, we turn to the second research question, which is about the influence from the family language on Italian in the domain of determiners. To answer this question divergent patterns in determiner productions among bilingual children are considered.

As we have already shown in the previous section no significant differences emerged between monolinguals and ITA-ARA and ITA-SPA bilingual children. For the porpouse of this section, it is worth recalling here the example (34), mentioned in section 1:

34) CHILD1SPA
la volpe vuole mangiare **la capra questa**
the fox wants to eat ***the goat this***

This is the only example of a demonstrative cooccurring with the definite article in the same NP, that could seem resulting from the influence of Spanish on Italian, that does not allow this type of construction. It should actually be said that, given the pragmatically marked nature of the construction of an NP with a definite article and a demonstrative in Spanish, along with its grounding to specific situations where the speaker intends to show psychological distance or contempt for the referent (see 2.2), it is difficult to confidently assert an influence of Spanish on Italian in the example (34). In other words, it does not seem evident that the example found corresponds to the typical Spanish usage. Additionally, it should be noted that similar constructions are infrequent in Spanish input, given their situation-specific nature. Example (34) is indeed an hapax in the corpus of ITA-SPA narratives in Italian, where no other divergent constructions can be attributed to cross-linguistic influence (CLI). The same holds true for Arabic-speaking bilingual children. In sum, no major influence of family language in these two groups can be found at the level of determiner choices.

At the other hand, as already shown in section 1, major differences are found in the ITA-CHI group, that differ from the other groups both in the use of demonstratives and in the frequency of article omissions. Recall, from Chapter 2.4, that Chinese expresses identifiability of the referent using

demonstratives (in anaphoric contexts) or bare nominals, while it mainly uses the numeral for 'one' (*yi*) for non-identifiability.

From the analyses presented in Section 1, a clear preference for the use of bare nouns in identifiable contexts in the Chinese-speaking bilingual group emerged (albeit bare nouns are also present in non-identifiable contexts).

Given the magnitude of this divergence from monolinguals, the question I address here is whether all the CHI-ITA children follow the same pattern or are some children in this group that follow a Chinese-like pattern (Id = bare; Nid = indef.) and others that follow an Italian-like pattern (Id = def.; Nid = indef.). An individual analysis was then conduct on this group. Children in ITA-CHI bilingual group were classified as to what was their dominant response in identifiable and non-identifiable contexts. Table 4 shows a count of ITA-CHI children according to their dominant pattern in identifiable and non-identifiable contexts:

		Non-Identifiable		
		Def	Indef	Bare
Identifiable	Def	1	2	
	Indef			
	Bare		5	2

Table 4 Number of Chinese-speaking children according to their dominant pattern in identifiable and non-identifiable contexts.

Most of the children in ITA-CHI group (7 out of 10) show contrast between the two contexts. However, among these children only 2 are convergent with the monolingual pattern, while 5 have a Chinese-like pattern, showing a clear influence from the family language.

Only 1 child generalizes the use of definites, resulting in adequate choices in identifiable contexts, but in egocentric errors in non-identifiable contexts. 2 children, on the other hand, show generalized difficulties with the use of Italian determiners, both definite and indefinite, mainly resorting to bare nominals in both contexts. It is worthy underlining that this classification is based on predominant responses, in fact, all children produce both definite and indefinite articles, showing that this category is present in their competence. Only one child (CHILD9CHI) never produces

definite articles (his predominant pattern is bare-bare in both contexts, but instances of demonstratives and indefinites are still present in his narratives).

As shown in section 1, the ITA-CHI group, with identifiable referents, also uses more demonstratives than the other groups. If we look at the subtypes of identifiable contexts in which demonstratives occur, we note that the vast majority of them are concentrated in anaphoric contexts (with negligible examples in other contexts; see the table in the Appendix C). This aligns with expectations based on Chinese grammar, which preferably employs demonstratives to refer back to entities mentioned earlier (anaphoric usage; see 2.4). Some examples of demonstratives in anaphoric contexts are given below (57-60), compared to determiners used in the same context (i.e., same referent in the same picture) by children from the other groups:

57) CHILD10ITA	58) CHILD5ARA	59) CHILD3CHI	60) CHILD4SPA:
la volpe saltò	e **la volpe** fece un salto	e **questa volpe** vuole mangiare	dopo **la volpe** aveva preso le pecore
the fox *jumped*	*and* ***the fox*** *make a jump*	*and* ***this fox*** *wants to eat*	*then* ***the fox*** *had taken the sheep*

The picture represents the moment when the fox decided to eat one of the two kids, jumped and caught it. Example (59) well represents the fact that Chinese-speaking children used the demonstrative where an Italian monolingual would prefer the definite article (see 57). It should be underlined that these are preferential, but not exclusive, trends. Actually, Italian allows the use of demonstratives in anaphoric contexts and examples of their use in monolingual children can be found in the corpus, such as the following:

61) CHILD3ITA
quella volpe lascià l'agnellino

that fox *left the little lamb*

Therefore the higher frequency of use of demonstratives in anaphoric contexts by ITA-CHI children seems to suggest a cross-linguistic influence of a qualitative nature, wherein a structure also used by monolinguals is overused by bilingual children (see 3.3). This is consistent with several studies, that found the use of demonstratives as an equivalent of definite articles in article-less languages, such as Mandarin, Russian, Malay, and Polish in contact with article languages (Polinsky, 2006; Moro, 2016; Otwinowska et al., 2020; Zhou et al. 2022).

Based on the predictions made in sections 2.5 and 3.5, we do not expect particular effects of Chinese on Italian concerning word order, given the strong convergence between the two languages (both SVO). Only the case of identifiable objects is a plausible area of influence, as to receive a definite reading in Chinese, an object must occupy the preverbal position, rather than the canonical post-verbal position that tends to raise an indefinite reading. However this position would lead to ungrammatical (or highly marked) sentences in Italian.

Looking at the data and considering subjects and objects together, non-identifiable referents (in first mention; n=44) are mostly encoded by post-verbal NPs (n=24), although the number of pre-verbal NPs (n=20) is only slightly lower (n=20). Identifiable referents (n=178) are mostly pre-verbal (n=97; all subjects), but with slightly fewer being post-verbal (n=81).

Separating the syntactic roles, of all the subjects produced (n=133), 77% (n=104) are preverbal and 23% (n=29) are post-verbal. However, preverbal subjects are not limited to identifiable referents; they mostly occur in identifiable contexts (n=97), but some instances are also present in non-identifiable contexts (n=20), with bare nouns and all the three kinds of determiners being present in both contexts, except for demonstratives in non identifiable contexts. As expected, the post-verbal subjects all occur in presentative sentences, introduced by *c'è* (‘there is’), or by inaccusative verbs (e.g., *tornare*, ‘come back’), which are perfectly grammatical in Italian and also expected in Chinese. However, it should be noted that in *c’è* constructions, NPs are not always indefinite (with bare nouns and all the three kinds of determiners being present, except for demonstratives; e.g. *c’è l’uccellino, ‘*there is the bird*’*, see also example 62 containing a bare noun), and that the referents of indefinite

NPs are not always new (discourse-integration errors, see 62 where the referent had already been introduced in previous discourse). Among the objects (n=60), only one case of preverbal object occurs in the corpus (see 63) with an identifiable referent (already introduced). The non-systematicity of the kind of occurrency presented in (63) does not allow us to define the case as an instance of CLI[69].

63) CHILD5CHI	63) CHILD10CHI
ha tornato **mucca piccolino**	**cane** l'uccellino prendere
*has come back **cow little***	***dog** the little bird takes*
***the little cow** has come back*	*the little bird takes **the dog***

It seems possible to say that Chinese-Italian bilinguals are not strongly guided by the definiteness-inclined (preverbal) or indefinite-inclined (postverbal) positions predicted by the grammar of their family language, but tather by adherence to the SVO order, which is common to both Italian and Chinese and that they preferibly mark the referent's identifiability status by using other devices (e.g., bare nouns/definite articles/demonstratives vs. indefinite articles as mentioned above). Furthermore, in presentational sentences, they are capable of subject-verb inversion, which is a point of convergence between the two languages, but they do not use these structures exclusively to introduce new referents. This seems consistent with the findings of Zhou et al. (2022). In a recent study on English-Chinese bilingual children (4-6 years), the authors found a preference for the pre-verbal position when introducing new referents, instead of the more appropriate VS inversion in Chinese. One possible explanation advanced by the authors is that the association between newness and post-verbal position was weakened by intense expoxure to English, which is strictly a subject-first language. The possibility of VS inversion present in both Chinese and Italian thus seems to facilitate the acquisition of this strategy, which, however, appears to be independent of the marking of identifiability. Further ad-hoc research to confirm these findings is necessary.

[69] It should also be noted that this type of word order exists in Italian, but it is highly pragmatically marked.

To summarize, no major effects from the family language are observed in the domain of determiners among bilingual children who speak Arabic or Spanish at home. It should be considered that the areas of divergence between the Arabic and the Italian article systems are concentrated in the non-referential NPs, which have not been included in this analysis and in other contexts in which Italian would use bare nouns while Arabic definite articles (e.g. plurals within the scope of the negation), which are not attested in the collected corpus. We do not know whether clear effects of the family language would have emerged in the areas of more marked divergence. In contrast, a more significant influence from Chinese is evident in the productions of ITA-CHI bilinguals. This influence is manifested in the overuse of demonstratives for anaphoric reference and the adoption of a Chinese-like pattern in distinguishing identifiable and non-identifiable referents: bare nouns for the former and an indefinite article for the latter.

3. Impact of experience and bilingual development

In this section we turn to the third research question, which is on whether bilingual language development and experience in Italian have an impact on divergences in their article system.

A description of patterns of development of the bilingual groups compared to monolinguals is given, followed by a comparison between monolingual and bilingual linguistic experience. Subsequently, the child-internal and child-external measures relevant for target-like uses of determiners in Italian are individuated.

3.1 Comparing the four groups: linguistic development and experience

Below is the comparison between the linguistic development of monolinguals and bilinguals, using the child-internal measures described in Chapter 4.

Recall here that the age range was 4;0 – 8;3 and that the groups were comparable in age (see participant description in 4.1). A Wilcoxon test was run to assess the age differences between the groups. The results indicate that there are no statistically significant differences in means among the groups: the p-values ranged between 0.2 and 0.3 in the comparisons between the monolingual group and each bilingual group, whereas they exceeded 0.6 in the comparisons between bilingual groups (see 4.1).

To verify the impact of age on target uses, a Kendall non parametric correlation was conducted between age and accuracy (ratio of target uses over the total NPs produced). Age and target uses were found no to be correlated, $\tau = 0.027, p = 0.80$. This is not in line with the expectations that target patterns should increase with age, nor with other studies on monolinguals (see Chondrogianni & Marinis, 2015 among others) or bilinguals (see Lindgren et al., 2022 among others). However, as noted by De Cat (2013:59) the frequency of errors in children (especially egocentric errors) shows considerable variability among groups of similar ages, not only across different studies but also within the same studies (see also Schultz & Grimm, 2019, see 3.4). Maratsos (1976) for example notes a significant amount of variation among the age groups in his study, and this variability was not directly linked to age, i.e., in his groups of participants some of the five-year-old children performed worse than the three-year-old children. Furthermore, it should be added that the present study did not take into account the age of first exposure to Italian in bilingual children, which plays a significant role in the development of linguistic competence (see 3.4). The internal structure of each group should also be considered, in which different ages are not equally represented.

As for child-internal factors, Working Memory (WM) and linguistic development measures are considered. In Table 5 decriptive statistics of WM scores are given.

	ITA			ITA-CHI			ITA-SPA			ITA-ARA		
(max)	**Range**	**Mean**	**sd**	**Range**	**Mean**	**sd**	**Range**	**Mean**	**sd**	**Range**	**Mean**	**sd**
WM (9)	1 to 8.5	6.50	2.66	2 to 8.5	5.60	2.15	2 to 7	5.25	1.90	0 to 8	5.3	2.58

Table 5 Overview of WM per language group.

Looking at Table 5, it si possible to note that all groups achieve an average of 5 or 6 correct answers (out of 9) in the working memory test, in line with the normative data of the instrument (Usai et al., 2017, see 4.2). In order to know whether there exist differences between the monolingual group and each bilingual group, a Wilcoxon test was run. No statistically significant differences in means of WM scores among the groups emerged: the p-values ranged between 0.06 and 0.4 in the comparisons between the monolingual group and each bilingual group. p-values exceeded 0.7 when the same test is run to compare bilingual groups to each other.

Turning to linguistic development, the measures considered in this study are reiterated here (see 4.4.1 for the description and computation procedure of each measure):

- Sentence Complexity: MLU5 and SubIndex
- Lexical Diversity: Vocabulary Diversity (VOCD), Number of different Verbs (NDV)

Descriptive statistics of these measures by language group is given in Table 6.

	ITA		ITA-CHI		ITA-SPA		ITA-ARA	
(max)	**Range**	**Mean (sd)**	**Range**	**Mean (sd)**	**Range**	**Mean (sd)**	**Range**	**Mean(sd)**
MLU5 (26.80)	14.8-26.8	21.45 (3.24)	9.25-17.8	14.26 (2.47)	17.6-26.6	22.25 (2.24)	14.4-24.2	18.36 (2.74)
SubIndex (1.34)	0.59-0.85	0.74 (0.08)	0.46-1.00	0.81 (0.19)	0.40-0.80	0.64 (0.15)	0.26-0.87	0.61 (0.18)
VOCD (46.02)	19.4-46.02	28.84 (7.54)	10.09-32.18	19.30 (7.30)	18.83-44.06	28.02 (7.28)	13.99-26.99	19.12 (3.75)
NDV (32)	17-32	21.45 (4.67)	5-17	12.30 (3.91)	12-28	20.10 (5.21)	12-26	17.3 (4.71)

Table 6 Linguistic profile in Italian, by language group.

Overall, the monolingual group appears to be more proficient than the bilingual groups. The differences seem to follow the degree of typological distance between the family languages and Italian, with the ITA-SPA group showing a linguistic development very similar to that of monolinguals, the ITA-ARA group being slightly more distant and the ITA-CHI group appearing less

proficient. To evaluate whether the differences with monolinguals are significant, a Wilcoxon test was conducted between the ITA group and each bilingual group, for each linguistic measure.

As for ITA-SPA group, results show no statistically significant differences between this group and the monolingual group for any measures ($p > 0.08$ for all the measures).

For the subordination index (Sub Index, i.e., the average number of clauses per sentence over the total sample per individual; Restrepo et al., 2010) the test did not show significant differences for any group ($p > 0.08$ in all groups). This is probably due to the structure of the task, whereby each child was asked to describe one figure at a time. The children, regardless of language group, usually described the pictures with one main clause and moved on to the second event using mostly "and" or "then", generally resulting in a list of main and coordinate clauses. The structure of the task itself was not suitable for eliciting more complex syntactic constructions.

As for the other measures the ITA-ARA and ITA-CHI groups are significantly different from the monolingual group: MLU5 (ITA-ARA $p = 0.01$; ITA-CHI $p < 0.001$), VOCD (ITA-ARA $p = 0.001$; ITA-CHI $p = 0.01$), NDV (ITA-ARA $p = 0.05$; ITA-CHI $p < 0.001$). In order to establish bilingual children dominance, a linguistic profile in the family languages was created[70] and a comparison between societal language skills and family language skills was conducted (see 4.4.1 for the description of the procedure for establishing dominance). The dominant language is the one with more skills with the highest-scoring.

As for dominance, in the Chinese-speaking group, 9 children were dominant in Chinese, and only 1 was Italian dominant. Among the Spanish-speaking group, 8 were found to be dominant in Italian and 2 were found to be dominant in Spanish. Among the children with Arabic as their family language, 9 were found to be dominant in Italian, only 1 was found to be dominant in Arabic. In sum, two groups are more dominant in Italian, namely ITA-SPA and ITA-ARA, while the ITA-CHI group is more dominant in Chinese. Dominance seems to be a variable that distinguishes the Chinese

[70] Since this chapter is devoted to Italian, the table with descriptive statistics of linguistic measures in the family languages is not reported here (see Appendix C).

speaking group from the other two bilingual groups and can therefore be considered to play a role in the accuracy of the choices of determiners, especially in determining omission and presence in obligatory contexts, which characterized this group.

We now turn to factors shaping children linguistic experience, which have been introduced in Chapter 4.4.3 and reiterated here:

- Input Quantity: Input sources (number of people speaking Italian to the child), Frequency of activities in Italian.
- Input Quality: number of linguistically relevant activities in Italian, parents' proficiency in Italian.
- Socio-economic Status: parents' level of education and occupation.
- Home literacy activities.

Descriptive statistics of child-external measures by language group is given in Table 7.

	ITA		ITA-CHI		ITA-SPA		ITA-ARA	
(max)	**Range**	**Mean (sd)**	**Range**	**Mean (sd)**	**Range**	**Mean (sd)**	**Range**	**Mean (sd)**
Input_sources (21)	18-21	20.73 (0,90)	6-18	10.00 (3.09)	6-17	11.60 (3.91)	8-15	11.38 (2.50)
Freq_activities (12)	12	12.00 (0)	3-11	6.90 (2.96)	7-12	10.50 (2.38)	7-12	10.57 (1.90)
Italian_activities (15)	15	15.00 (0)	6-13	7.90 (2.18)	9-11	11.00 (1.63)	9-15	12.57(1.81)
Nativ_parents (32)	32	32.00 (0)	8-24	16.70 (5.29)	20-32	24.80 (4.76)	14-25	19.33 (4.12)
Mother_Edu (4)	3-4	3.75 (0.5)	1- 3	2.75 (0.70)	3-4	3.25 (0.5)	2-4	3.25 (0.70)
Mother_Occ (5)	3-5	4.00 (1.15)	1-3	2.00 (1.00)	2	2.00 (0)	1	1.00 (0)
Father_Edu (4)	2-5	4.00 (0)	0-3	2.50 (0.97)	2-3	2.75 (0.5)	2-4	2.83 (0.75)
Father_Occ(5)	4	3.50 (1.73)	0-3	2.22 (1.09)	2	2.00 (0)	2 -3	2.33 (0.57)
Home_Literacy (104)	89-97	93.25 (3.30)	64-95	80.33 (11.35)	68-103	89.75 (15.64)	50-84	69.75 (12.71)

Table 7 Overview of child-external measures, per language group

Looking at Table 7 differences between linguistic experience of monolinguals and bilinguals are evident. A Wilcoxon test was run to assess whether these differences are significant, and the results indicate that indeed they are, for each measure for each group ($p < 0.03$ for each measure for each group) except for the mother level of education of ITA-ARA ($p = 0.21$) and ITA-SPA ($p = 0.18$) groups and for home literacy activities of ITA-SPA group ($p = 0.88$). This tells us that mothers in these two language groups have comparable levels of education to the mothers of the Italian children in this study, but with different levels of occupation[71]; furthermore, the results indicate that ITA-SPA children are involved in domestic early literacy practices similar to those of their monolingual peers.

When bilingual groups are compared to each other, by running the same test, differences between ITA-SPA and ITA-ARA never reached the level of significance ($p < 0.06$ for each measure), showing that children in these two groups undergo a highly similar language experience. More different appear to be the Chinese-speaking group. The results of Wilcoxon test indicate that there are no significant differences in number of people speaking Italian to the children in Chinese communities and in the other two speech-communities ($p > 0.2$ for both the comparison to ITA-ARA and ITA-SPA groups). Also home-literacy activities are not an area of divergence between the groups ($p > 0.1$ for both the comparison to ITA-ARA and ITA-SPA groups). Linguistically relevant activities, however, is an area of wide divergence between the Chinese-speaking group and the other two bilingual groups: test results show that ITA-CHI children carry out more activities in Chinese than in Italian (Italian activities: $p < 0.01$ for both ITA-ARA and ITA-SPA comparisons) and those in Italian with a significantly lower frequency than the other groups (Frequency of Italian activities: $p < 0.03$ for the comparisons to both the other bilingual groups). As for parents'proficiency in Italian, the results of the comparison show that even in this area the difference with the other two groups is significant ($p < 0.03$ for the comparisons to both the other bilingual groups), that is, Chinese parents are generally less competent in Italian than both Arabic-speaking and Spanish-speaking parents and

[71] For the occupation-education mismatch of immigrants in Europe see Aleksynska and Tritah (2014).

this can affect the quality of the input children receive. Overall, given that all children receive the same quality and quantity of input in Italian at school, the characteristics of the input to which Chinese children are exposed outside of the school are significantly different from those of other language groups.

As for socio-economic status no significant differences between the bilingual children families emerged from Wicoxon tests ($p > 0.1$ for both parents'level of education and occupation). Only the occupation of Spanish-speaking fathers seems to be an area of significant divergence (compared to ITA-ARA fathers: $p = 0.03$; compared to ITA-CHI fathers: $p = 0.006$). Caution is necessary in interpreting this result, given that information on fathers is the most missing information in the questionnaires, especially those from Spanish-speaking families. No difference in socio-economic status seems to be the most cautious and acceptable statement for this study.

3.2 Relevant factors for target uses of articles

After describing both the linguistic development of bilingual children and their linguistic experience, I asked which measure is connected with accuracy in the use of determiners.

It has already been shown that bilingual groups are characterized by different errors in the use of determiners in Italian: in ITA-ARA and ITA-SPA bilingual groups the most common error is the use of the definite article in place of the indefinite article (egocentric error), that also occurs in ITA-CHI group. The use of indefinites seems to be an area in which the distance with monolinguals emerges, since they produce a higher number of indefinites in non-identifiable contexts. However, the ITA-CHI group compared to the others mainly commits errors of omission, especially in contexts of identifiability, resulting in a large quantity of ungrammatical NPs.

In the following subsections the two types of errors are considered separately, as substitution errors are linked to discourse-pragmatic factors, while omission errors are grammaticality errors, linked to syntactic-semantic constraints.

3.2.1 Relevant factors for discourse-pragmatic accuracy of using determiners

The first type of errors will be considered here. The level of accuracy of the determiners, when produced, was calculated for each child, i.e., the ratio between the number of target determiners (definites/demonstratives in identifiable contexts and indefinites in non-idenifiable contexts) over the total of articles present in the individual sample. In order to see which child-internal and child-external measures are most relevant for the target production of articles in bilingual children, scatterplots were created and non-parametric Kendall correlations conducted (considering the bilingual group as a whole).

Only two measures of language development were found to be relevant: MLU5 and NDV (see Figures 5 and 6).

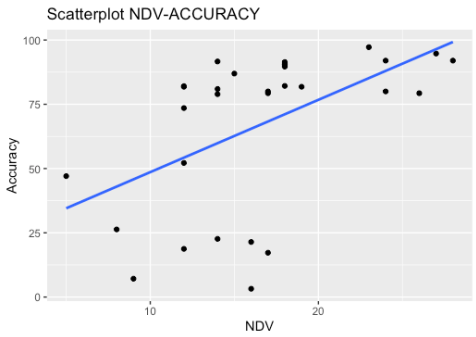


Figure 5 Scatterplot MLU5-ACCURACY in the use of determiners by bilinguals

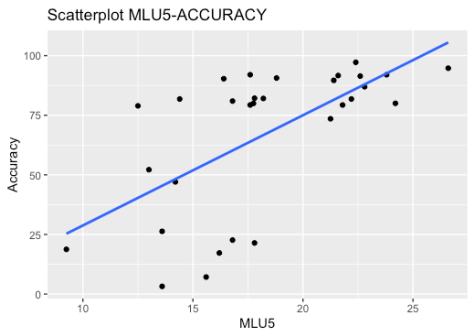


Figure 6 Scatterplot NDV-ACCURACY in the use of determiners by bilinguals

Both the measures show a moderate correlation with the level of discourse-pragmatic accuracy (MLU5-ACCURACY: $\tau = 0.48$, $p < 0.001$; NDV-ACCURACY: $\tau = 0.44$ $p < 0.001$)[72]. From Figure 5 it emerges that children with a lower sentence complexity make more errors with determiners (which generally is the use of definites in non-identifiable contexts). The same can be said for the breadth of the vocabulary, at least the verbal one (NDV). In the type of narrative task used in the

[72] When adjusted for Bonferroni correction for multiple comparisons ($p < 0.01$) those two measures remained significant.

present study, the pictures showed few nominal referents, so it seemed that verbal vocabulary could better capture differences in lexical breadth. Children who make less substitution errors are those with a larger vocabulary. This seems to confirm that there is an interplay between morphosyntactic and vocabulary development and children's pragmatic skills in the acquisition of determiners (Rozendaal, 2008).

Various considerations can be made regarding the lack of correlation with WM. First of all, the low quantity of data analyzed and participants in the experiment must always be taken into consideration. Furthermore, the groups involved a rather wide age span, during which cognitive development is characterized by strong changes with respect to executive functions and WM (Garon et al., 2008). Some ages were underrepresented within the groups, and these developmental leaps could be barely visible in the sample. This may have led to a mixed picture, which did not reveal clear WM effects. It is also possible that the presentation of one picture at a time lightened the demand on working memory. These considerations taken together lead us not to take this result as an indication that WM does not play a role in article choice at all.

As regards the external variables, those that have proven to be most relevant for the target-like use of the determiners are the linguistically relevant activities in Italian, both in terms of frequency and quality of the input. See figures 7 and 8 below:

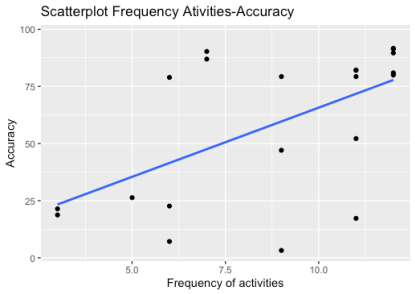


Figure.7 Scatterplot Frequency of activities in Italian-ACCURACY (%) in the use of articles in bilinguals

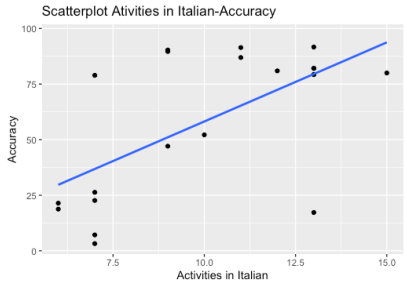


Figure.8 Scatterplot Number of Activities in Italian-ACCURACY (%) in the use of articles in bilinguals

Both external measures show a moderate correlation with the level of discourse-pragmatic accuracy (FREQ_ACTIVITIES-ACCURACY $\tau = 0.46$ $p = 0.005$ e ITALIAN_ACTIVITIES-ACCURACY $\tau = 0.39$ $p = 0.01$). When adjusted for Bonferroni correction for multiple comparisons ($p < 0.005$) only the frequency of activities in Italian remained significant. This suggests that input from multimedia sources, from gaming activities alone or with classmates in Italian are particularly relevant for acquiring good levels of sensitivity to contextual factors (e.g., shared knowledge and common ground) and receiving rich and diverse input.

3.2.2 Relevant factors in article omissions: the CHI-ITA group

Considering only the Chinese-speaking group, a distinct type of error can be separated, specifically related to grammaticality — the omission of determiners in obligatory contexts, a characteristic trait of this group. To discern which measures of linguistic competence exert a more significant impact on this target-deviant pattern, I present descriptive statistics on the linguistic development of ITA-CHI children. They are categorized based on the dominant pattern in determiner production in the two contexts (identifiable and non-identifiable): bare-bare, bare-indef, def-indef (the child with a def-def response-pattern was placed in the last subgroup as he does not produce ungrammatical responses). Refer to Table 8 below for details:

	Bare-Bare (2 children)		Bare-Indef (5 children)		Def-Art (3 children)	
(max)	**Range**	**Mean (sd)**	**Range**	**Mean (sd)**	**Range**	**Mean (sd)**
MLU5 (17.8)	13.6-16.20	14.90 (1.63)	9.25-17.80	14.61 (3.38)	12.5-14.20	13.23 (0.87)
SubIndex (1.00)	0.46-0.53	0.49 (0.05)	0.73-0.98	0.85 (0.13)	0.94-1.00	0.97 (0.03)
VOCD (32.18)	13.21-20.71	16.96 (5.30)	11.41-32.18	20.46 (7.72)	10.09-29.59	18.90 (9.88)
NDV (17)	16-17	16.50 (0.70)	8-16	11.80 (3.34)	5-14	10.33 (4.72)

Table 8 Overview of linguistic measures of ITA-CHI children by response-pattern group

In interpreting these results, the size of the samples must always be kept in mind. Furthermore, the high variability in the command of Italian in this group of bilinguals should be noted. As shown in the previous section, for the ITA-CHI group, Italian, the societal language, is not the dominant language and the one in which individuals receive quantitatively and qualitatively very different input from child to child (see Table 8 above). Linguistic skills also show high levels of individual variability. However, it can be noted that the measure of syntactic complexity (SubIndex) has the lowest mean in the group with a bare-bare response, a little higher in the group with a Bare-Indef response and even higher in the group producing NP with determiners as the main response in both contexts. Considering that the high number of bare nouns can be attributed to an influence from the family language (CLI), it is possible that children with a greater mastery of Italian syntax are better able to inhibit the option coming from the dominant language (Chinese) in the construction of the nominal phrase. The proceduralization of grammatical knowledge might be more efficient and faster in those children who can rely more consistently on the syntactic options for reference available in the non-dominant language (see Torregrossa et al., 2018). Observing table 8, even the breadth of vocabulary, measured through VOCD, seems to distinguish the bare-bare group from the other two groups. It therefore seems possible to align with the considerations made by Hervé and Serratrice (2018:783) who state that CLI would be, at least partly, determined by the children's overall level of expressive abilities, i.e., children with higher expressive skills are less prone to CLI.

Turning to child-external measures, the only measure that has proven relevant is related to the number of linguistically stimulating activities carried out in Italian. (ITALIAN_ACTIVITIES-OMISSION_RATE: $\tau = -0.26$ $p = 0.33$) . See figure 9 below:

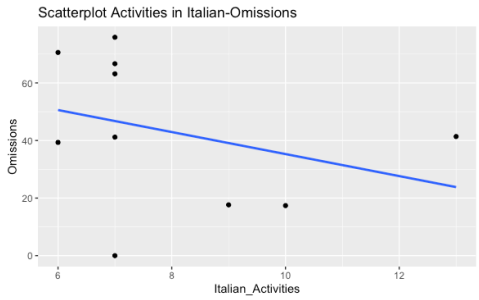


Figure 9 Scatterplot ACTIVITIES IN ITALIAN-ARTICLE OMISSIONS RATE (%) in bilinguals

As in the previous case, also for the ITA-CHI group, the external measure that is most relevant for a low percentage of article omissions (calculated as the ratio of ungrammatical bare nouns to the total number of NPs, converted to a percentage) is the input coming from activities such as playing with multimedial devices. These provide a rich, diverse, and authentic input and offer opportunities for linguistic exposure during engaging activities (Sun & Yin, 2020).

In summary, the linguistic development and linguistic experience of the three bilingual groups appear different from that of monolinguals and differ from each other. While children in the ITA-ARA and ITA-SPA groups are generally dominant in the societal language, the ITA-CHI group appears more dominant in the family language. The linguistic command of this bilingual group in Italian is the furthest from both that of monolinguals and of other bilingual groups. The linguistic development of the ITA-ARA and ITA-SPA children is closer to that of monolinguals. When compared to each other, the differences between these two bilingual groups did not prove to be significant. There are no differences in the socio-economic status of the families involved in this study. The linguistic experience of the ITA-ARA and ITA-SPA groups is very similar, whereas that of the ITA-CHI children appears completely different. The differences do not lie so much in the

quantity of input they receive from people in their community but in the quality of this input (measured through the proficiency of the parents) and the number and frequency of linguistically relevant activities conducted in Italian. The ITA-CHI group appears to be the most disadvantaged in Italian in this sense.

In seeking which linguistic and non-linguistic factors were relevant for the target use of articles, it was found that involvement in linguistically relevant activities appears important for both the production of grammatical NPs (i.e., with determiners in obligatory contexts) and for pragmatically target use based (i.e., definites in identifiable contexts, indefinites in non-identifiable contexts). It was then observed that children with greater sentence complexity and a broader vocabulary tend to produce more contextually appropriate articles, and that both syntactic skills and vocabulary breadth are relevant for the production of determiners in obligatory contexts.

Finally, it should be pointed out that for the analysis, one variable at a time has been considered. However, the results may be different if all variables were considered in a unique model (e.g., a multiple logistic regression). Given the limited sample size, there exists uncertainty regarding whether the sample meets the necessary requirements for a reliable logistic regression analysis. Moreover, logistic regression assumptions, including the absence of multicollinearity among independent variables, linearity of the logit, and distribution of residuals, cannot be assured in the sample of the present study. Hence, to mitigate potential risks associated with violating these assumptions, a correlation analysis considering a variable at a time has been preferred.

4. Summary and Discussion

This section is dedicated to summarizing the results presented in the previous sections and their discussion in relation to the previous literature. Let us recall the research questions and the corresponding expectations formulated in chapter 3.5:

RQ1. Do bilinguals and monolinguals distribute articles differently according to identifiability?

Recall that the children are 4-8 years old. Based on the facts that at that age, the cognitive skills underlying pragmatic abilities for reference are already in place (Rozendaal, 2008) and that the referential systems of Italian and the involved family languages differ to a varying extent, I expect that all children have already developed sensitivity to the distinction of identifiability, but at the same time, the distribution of determiners would be more similar to the one found in monolingual for those individuals whose family language more closely resembles Italian with respect to its referential system. Considering that children were administered a narrative task, cognitively more demanding than spontaneous conversations, it is expected that some errors may emerge. It is expected that the most common error in all groups would be the use of the definite article in non-identifiable contexts (egocentric error).

RQ2. Does the family language influence the societal language in the acquisition of definiteness?

The expectation is that Cross-Linguistic Influence (CLI) from the family language to Italian would manifest in the domain of determiners if there was partial overlap between the two languages:

- for the Spanish-speaking group, a visible effect of the family language on Italian is not expected, given the total convergence of the article systems in the two languages, with respect to the phenomena under investigation. The only area of divergence where the effects of Spanish on Italian can be expected is the co-occurrence of the article and the demonstrative in the same NP.

- As for the Arabic-speaking group, the article systems of spoken varieties of Arabic and Italian substantially converge, with respect to the phenomena under investigation. Areas where the effects of the family language might be evident include the co-occurrence of the article and the demonstrative in the same NP in anaphoric contexts. It is also expected that there would be less extensive use of bare nouns by Italian-Arabic bilingual children compared to monolingual peers, in

contexts where Arabic prefers the use of the definite article while Italian allows bare NPs: mass nouns in PPs and bare plurals in non-identifiable contexts.

- For the Chinese-speaking group, an effect of Chinese is expected in two forms:
    - overuse of demonstratives in an anaphoric contexts,
    - omission of the article in identifiable contexts.
    - preverbal object with identifiable referents

RQ3. Is there an effect of child-external and internal factors on the acquisition of definiteness? Given the interplay between morphosyntax and pragmatics, it is expected that children with greater dominance in Italian morphosyntax would also have higher levels of accuracy at the discourse-pragmatic level (fewer substitution errors) and less article-omissions. It is also expected that greater quantity and better quality of exposure would be linked to more target-like use of articles (fewer omission or substitution errors).

4.1 Summary

In this study, I aimed to analyze the expression of definiteness in a narrative task in Italian administered to bilingual and monolingual children (aged 4;0-8;3 years). Three groups of bilingual children were identified, all with Italian as the societal language and three different family languages: Spanish, Arabic, and Chinese. A comparison was made both among the three bilingual groups and with the Italian monolingual control group, matched for age.

The results of the analyses revealed that bilingual children, like monolinguals, are sensitive to the identifiability status of the referent and know the main coding strategies required in Italian. When the referent is identifiable, all groups produce more NPs with the definite article, apart from the group of Italian-Chinese speakers who show a strong preference for bare nouns. The Chinese-speaking group is also the one that makes the most extensive use of demonstratives compared to the others, but

this use is limited to anaphoric contexts. This group also produces a slightly higher number of indefinites in identifiable contexts. When the referent is non-identifiable, all groups produce more NPs with indefinite articles, confirming their sensitivity to contextual factors. Overuse of definites instead of indefinites is the most frequent errors in this type of contexts and it is more pronounced in bilinguals than in monolinguals. Again, the Chinese-speaking group shows a higher frequency of bare nouns in this type of contexts. With non-identifiable referents monolinguals and Spanish-speaking bilinguals do not produce bare nouns at all, while Arabic-speaking children produce very few instances of bare nouns, but all attributable to a single child. The expectations formulated for this research question seem to be met.

Patterns of divergence shows effects of family language only in the ITA-CHI group. Cross-linguistic influence manifests itself both in the overuse of demonstratives in anaphoric contexts compared to monolinguals, and in the Chinese-like coding pattern of the identifiability contrast: bare nouns for identifiable referents and indefinite articles for non-identifiable referents. The first case shows an influence that reflects the preferential tendencies of Chinese which in anaphoric contexts uses demonstratives, where languages like Italian would use definite articles. In the second case, the outcome of the CLI is ungrammatical in Italian with omissions not attested in the narratives of Italian monolingual peers. As long as ITA-CHI children are concerned the expectations formulated for the second research question seem to be met. No major effects of Spanish and Arabic have been found in the narrative productions of the other two bilingual groups.

Considering bilingual children together, results of analysis also show that target article use, in terms of contextual accuracy, was higher in children with higher MLU5 and richer verbal vocabulary, but other important variables like, SubIndex and VOCD were not. However, the SubIndex and VOCD were important in determining the quantity of ungrammatical bare nouns in the Chinese-speaking group and seem to be therefore more related to grammaticality errors. Among the external variables, those linked to linguistically relevant activities (such as input from multimedia sources, games with classmates, etc.) were found to be particularly important for a target-like use of the article, both in

terms of discourse-pragmatic accuracy and grammaticality. Neither age nor working memory were related to accuracy in the use of articles.

4.2 Dicussion

The first aim of the study was to verify if the distribution of determiners based on the identifiability of the referent in the three groups of bilingual children and monolinguals was the same. The results of the study confirm that there are no differences in contextual sensitivity among the groups, all of which seem sensitive to the abstract dimension of referent identifiability, coded, however, with different formal means. The results are therefore in line with studies that have found early pragmatic sensitivity (Skarabela & Allen, 2013; De Cat, 2011, 2013). Identifiability is a language-independent category, to which bilingual children in the study appear to have already developed sensitivity at the age of 4. Indeed, errors of discourse-pragmatic adequacy (substitution errors) are few, and most are egocentric (use of definites instead of indefinites).

As it emerged from previous literature, adequacy errors can be linked to cognitive factors and, according to general speech production models, are generated at the conceptual level, where the speaker decides on the informational content of the linguistic message to be communicated, as related to the amount of knowledge shared with the interlocutors: whether referents are identifiable or not to the hearer (De Bot, 1992; De Lange, 2009; Levelt, 1989; see chapter 3.1.2) . Narrative tasks require a large processing demand, and it is possible that the pictures contributed to errors in adequately integrating perceptual and discursive cues. In fact, in order to lighten the cognitive load during the task, children tend to rely a lot on the visual context, to presuppose joint attention and common ground, even in the absence of shared figures between child and listener and this would lead to evaluations of referent identifiability status that are not in line with the knowledge state of the interlocutor (De Cat, 2013; Serratrice & De Cat, 2020). This effect may have been exacerbated by the fact that the scenes were presented one by one. This type of presentation, in fact, may have

compromised the evaluation of the cues from the previous discourse and the maintenance/updating of the discourse model. Children would concentrate solely on the target picture, without allocating resources on representing the preceding event and the coherence relation between events and referents. For bilinguals, these demands are further compounded by the inhibition operations from the non-target language, during formal coding. This would explain the slightly higher number of egocentric errors in bilinguals compared to monolinguals.

Keeping track of the reference during the unfolding discourse is indeed a complex task, which requires the activation of the executive functions for the management of one's own speech model (activation of the referent, updating and adaptation of the model to the listener's knowledge) and at the linguistic level it requires the child to know how to use articles also as elements of textual cohesion. Doing this while turning one page at a time may have made the task even more difficult.

The bilingual groups that are closest to monolinguals are ITA-ARA and ITA-SPA groups, which are extremely similar to each other. These results are in line with Zdorenko and Paradis (2012), showing that Arabic-English bilinguals don't differ from the Spanish-English bilingual children in using English articles. However, in their work the authors started from the description of Modern Standard Arabic, which led them to state the absence of an indefinite article in Arabic. Actually, Modern Standard Arabic is not the language in use in everyday interactions among Arabic speakers. And spoken varieties vary from the acrolect also in the domain of definiteness, precisely due to the presence of an overt element (e.g., the numeral 'one') used as an indefinite article (Mion, 2020:187). It is not surprising, therefore, that Arabic speakers and Spanish speakers exhibit similar patterns, both resembling monolinguals. The authors argued that the surprising lack of indefinite article omissions with non-identifiable referents in the Arabic-speaking group is due to the fact that the specific transfer of the [– definite] feature mapping did not occur in these children. This seems no to be the case, and a less complex, more parsimonious explanation would be preferable. In non-identifiable contexts, Arabic and Italian converge in the use of an indefinite article and therefore the options provided by the two languages are not in competition.

In addition to errors of pragmatic adequacy, bilingual narratives also feature grammaticality errors, i.e., the omission of determiners in obligatory contexts, that does not occur in monolingual productions. While in the groups of Arabic and Spanish speakers these errors are almost entirely concentrated in the productions of a single individual per group, the omission error is the most frequent and widespread among Chinese-speaking children.

With reference to speech-production models, while substitution errors are attributable to the so-called conceptual level, errors of determiner omission are assignable to the lemma level, the level of linguistic decision-making: grammatical encoding and lemma selection (De Lange et al., 2009).

In the group of Chinese-speaking children the frequency of omissions is very high and seems to be attributable to cross- linguistic influence (CLI) from the family language. Indeed, many children appear to follow a Chinese-like pattern in distinguishing identifiable from non-identifiable referents: bare nouns for the former, indefinites for the latter. The more frequent use of demonstratives is also an effect of family language, which appears only in anaphoric contexts, exactly like in Chinese (Jenks, 2018). In anaphoric contexts, Chinese uses demonstratives where languages such as Italian or English would use the definite articles. Since the use of demonstratives for anaphoric reference is a strategy also employed by monolinguals, it can be argued that this is a case of cross-linguistic influence manifesting itself in terms of overuse of a grammatical structure.

These results are consistent with various accounts. The data presented in the study can be explained both by the model proposed by Müller and Hulk (2001), with Chinese being the simpler system thus affecting Italian in a phenomenon related to the syntax-pragmatics interface, and by processing accounts (both shared-syntax and connected-syntax; see chapter 3.3.2). However, since Müller and Hulk's proposal has encountered several counterexamples and processing-based models seem to better account for the available data in many studies (including Hervé and Serratrice's 2018 study on determiners in bilinguals), I lean towards an explanation based on processing and speech production models, favoring the shared-syntax account. While the results are compatible with both shared and connected-syntax accounts, the shared-syntax model will be preferred here as it is the only

one with an adapted developmental version. The connected-syntax account, in contrast, has not proposed a developmental model for the formation of indirect connections between representations, leaving it unclear how these connections develop (Unsworth, 2023).

Below, a speech-production model (with shared syntax at the lemma level) is presented, as adapted from Arnold and Zerkle (2019) and Unsworth (2023) to describe the case of the production of the Italian NP for the referent "goat" by an Italian-Chinese bilingual child (Figure 10):

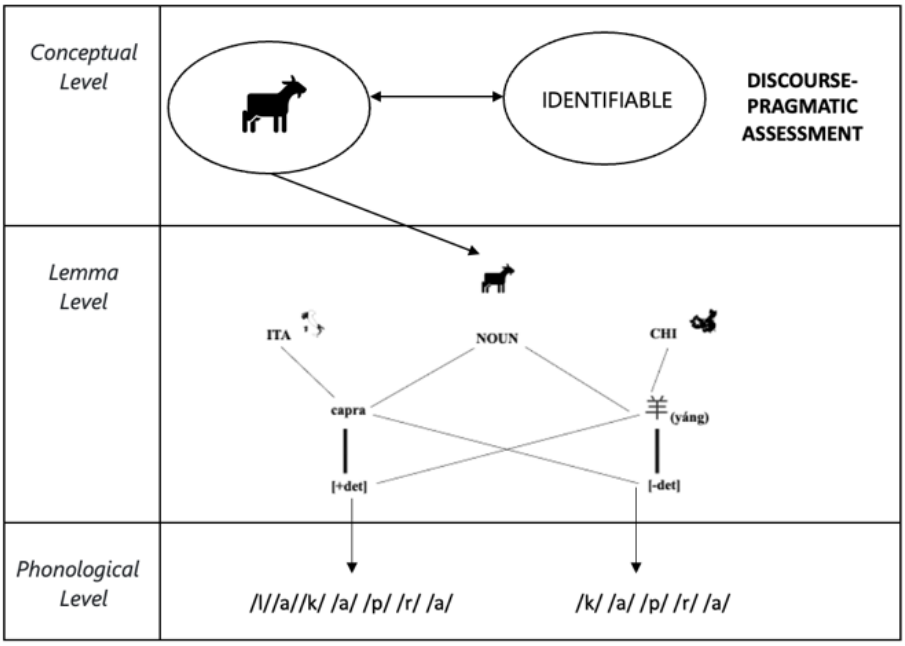


Figure 10 Speech production model for bilinguals (Italian-Chinese)

Once the evaluations concerning the identifiable status of the referent have been made at the conceptual level, and the accessibility index is assigned (3.1.2), the child has to make formal choices at the lemma level. In constructing the surface structure of the noun phrase (NP), the grammatical encoder must decide whether to include an article or not, and if so, which article to select to better indicate the identifiability status of the referent (as indicated by the accessibility index; see 3.1.2). However, at the lemma level, the syntactic representations of the two languages are shared and can enter into competition.

With the schema in Figure 10, it is not intended to imply that the determiner is part of the lemma 'goat'. In the preverbal message originating at the conceptual level, there is not yet a separate encoding for the semantic information of the article; simply, the referent is assigned an accessibility index indicating whether it is identifiable to the interlocutor. Once the verbal message moves to the formulator, the corresponding lexical lemma for the referent is activated, which contains a series of semantic and syntactic information, including the syntactic category of the lemma, noun in the example. If the noun functions as the head of the phrase, the grammatical encoder (see 3.1.2) initiates a series of operations for constructing the NP, including the selection of a determiner, according to language-specific rules that in bilinguals come into competition. The schema is a simplification and aims to account for the lexical-based nature of this mental process, where the noun is the driving force for activating the rest of the lemmas necessary to form the NP. The lines connecting the noun to the determiner represent possible competing activation paths.

In the case of the non-identifiable referent, Chinese uses the numeral *yi* ('one') (which has almost completed its grammaticalization path and in many cases acts as an indefinite article, see chapter 2.4) and Italian does the same with the indefinite article, which has the same form as the numeral 'one'. Therefore, in this case the two languages rarely enter competition. If, at the conceptual level, the referent is evaluated as identifiable, at the lemma level the options "NP with the determiner" ([+det]) from Italian and "bare noun" ([-det]) from Chinese, associated with the selected lemma ('goat' in the example), come into competition. If the evaluation of the referent leads to the recognition that it is identifiable on an anaphoric base, then the realization of [+det] can also be a demonstrative, which is an option available for both languages.

Several factors can contribute to the prevalence of the [-det] option on the other:

- **priming effect:** the [-det] option could be activated either by listening/using bare -preferring Chinese structures or from listening/using structures in L2 Italian of Chinese learners, which is often characterized by omissions of the determiner (Chiapedi, 2010). Once this option is

activated, its residual activation makes it more readily accessible to the speaker also in the successive productions (Unsworth, 2023).

- **inherited frequency:** the Italian NP structure can inherit the very high-frequency of the Chinese NP structure without determiners.
- **dominance effect:** it was seen that the group of children speaking Italian and Chinese is strongly dominant in Chinese. The high occurrence of omissions and few occurrences of definite articles is in line with the result of Van Dijk et alii (2021) according to which dominance would determine the strength of the CLI. However, it should be added that in our case dominance was not calculated based on exposure to the societal language, as in Van Dijk et alii (2021). Especially in the case of Chinese-speaking children in Italy, the equivalence societal language = language to which children are most exposed = dominant language does not hold. In fact, if one look at the linguistic profiles of the three groups, it emerges that the Chinese group is strongly dominant towards the family language, to which it is much more exposed, despite the children being all included in the Italian school context.

Examining the overall proficiency of the Chinese-speaking group, it becomes evident that these children have a lower proficiency in Italian compared to their family language and to the other groups. The influence of overall expressive abilities on the probability of CLI seems to align with the expectations outlined by Hervé & Serratrice (2018). According to the authors, CLI would be, at least partly, determined by the children's overall level of expressive abilities. Children with higher expressive skills are less prone to CLI.

Although rare, cases of omission are also present in Arabic and Spanish speakers, but they are mostly produced by two individuals. A more specific analysis is not possible, since we don't have information about speech impairments of these two children, a condition that can lead to article omissions (Belletti & Guasti, 2015; Marini et al., 2008). An alternative possible explanation can state that omission could be due to processing limitation, according to De Lange et al. (2009)'s account.

They argued that the more processing resources and time are necessary during the production, the more likely the article will be omitted. The girl in the Spanish group is the youngest of the group and is 4 years old, while the one in the Arabic-speaking group is an early L2 bilingual, having been exposed to Italian at the age of 6. It is therefore possible that these two girls have more limited morphosyntactic resources and that during the production they choose the cognitively lighter path, given that the narrative task is already cognitively demanding in itself. De Lange et alii (2009) also argue that the minimization of structure (e.g., article omission) can be explained in terms of optimization of the underlying processing mechanisms. The choice of the bare noun would appear the lightest from a cognitive point of view, since it avoids various operations, including morphosyntactic agreement between noun and determiner. If the morphosyntactic system in the target language is not fully developed, as in a very young simultaneous bilingual or in an early L2 child, then bare nouns are the easiest rout (De Lange et alii, 2009: 1530). Other non-linguistic factors, such as children tiredness or other unknown circumstances, may influence processing during article production.

From what has emerged so far, it seems possible to align with the conclusions of Rozendaal (2008), who suggested an interplay between morphosyntactic competencies and pragmatic skills in the selection of determiners. Pragmatic adequacy of determiners (which means knowing how to introduce and maintain the referents already introduced, or in the common ground) grows with the linguistic development of the child, both in sentence complexity and in vocabulary. The more competence increases, the more means one has to adequately integrate the referents into the discourse.

Quality and quantity of input also play an important role. This is in line with the considerations made by Schulz and Grimm (2019:15) according to which the more input learners receive in the target-language the faster they can make the necessary language-specific choices. This would also explain the slightly higher number of substitution errors in identifiable contexts made by Chinese-Italian bilinguals, who are less exposed and less dominant in Italian and thus have less resources to adequately coded the referent as identifiable to the hearer. Results also showed they have greater

command of indefinite articles than definites and this could also lead them to substitutions, i.e., using indefinites is a linguistically much easier, cognitively lighter, thus faster way to produce NPs.

In line with Torregorssa et alii (2021:703), it seems possible to argue that bilinguals with dominant experience in the non-target language, as the children in ITA-CHI group, are more unlikely to have developed a fast and efficient proceduralization of grammatical knowledge than children who are more balanced or dominant in the target language. Thus, Chinese speaking children would be less able to consistently rely on the syntactic options for reference available in the target language.

Last, according to Torregrossa et alii (2018: 19) the production of the syntactic options for reference available in the target language can be enhanced by greater and I would also add better exposure. Results in this study support their view, since for the use of determiners in obligatory contexts, i.e., the production of grammatical NPs, a fundamental role is played by external variables, in particular activities in Italian. Linguistically stimulating activities in which the child is involved provide a very rich and diverse input, where the child can encounter various and correct uses of determiners. Since the article system is an area of particular difficulty for adult L2 learners, especially for speakers of Chinese who often omit them, input from non-native adult speakers, might provide less evidence of appropriate article use. Additionally, activities in Italian often involve the child in production, so it is likely that the effect of these activities is also determined by opportunities for output, which were not measured in this study. Moreover, Valian (2015) states that engaging and enriching activities contribute to the development of executive functions, which constitute the cognitive underpinnings for referential choices (De Cat, 2015).

The conclusions reached in this study suggest that variation in the production of accurate (in) definites in the societal language among bilingual speakers seem to be motivated more by variables related to language competence than by variation in working memory (WM). This finding has implications for our understanding of the interaction between linguistic and cognitive variables in reference production among bilingual children. These factors jointly influence bilingual reference production,

together with an effect of the task, but likely interacting differently and with varying weights for different referential expressions (full NPs, pronouns, zero anaphora, etc.)

It appears that for the production of felicitous and grammatical (in)definite NPs, linguistic competence plays a particularly important role. This aligns with other studies on referential expressions in bilinguals, which have highlighted that the use of indefinite nominals in contexts of introducing referents improved as a function of language proficiency in at least one language (Zhou et al., 2022). Most studies that have addressed reference have emphasized the importance of working memory (WM) in making appropriate referential choices: it is related to storing information about the referent, and thus, limitations in WM may lead to difficulty in computing discourse and, more specifically, remembering the information related to a certain discourse referent, such as the grammatical role of its previous mention (i.e., antecedent). This is why variations in WM are often related to variations in referential choices, namely the tendency to under- or over-specify referential expressions. However, in the present study, the over- or under-specification of referential expressions, was not considered, since only (in)definite descriptions (definite or indefinite full NPs) were under investigation. It is possible that the choice between definite and indefinite determiners in a full NP, rely more on processes of retrieving and updating referents, based on other executive functions, than WM.

Furthermore, linguistic competence increases the speed of information processing, which has consequences for the accuracy of referential expression (Hendriks, 2016). This is because it allows the speaker to access forms more quickly and with less cognitive effort, while simultaneously considering the perspective of the listener, including the necessary inhibition operations of competing options (such as structures activated in the other language or competing referents in the discourse).

The task of making referential choices during narratives requires significant cognitive capacity, so children with greater linguistic competence are at an advantage.

6. Teaching proposal

1. Introduction

In this chapter, a teaching proposal on narrative competence will be described. It was directed towards teachers and students from the schools that took part in the research project outlined in the previous chapters. This educational action has emerged as a way of turning the favor to the community, in line with Wolfram's principle of gratuity (1993). The principle is articulated below (Wolfram, 1993:227):

"*Investigators who have obtained linguistic data from members of a speech community should actively pursue positive ways in which they can return linguistic favors to the community.*"

This principle connects to two further principles, previously proposed by Labov (1982), which can motivate linguists to take social action: the *principle of error correction* and the *principle of debt incurred*. According to the first principle, linguists should take action when a linguistically erroneous idea becomes widespread in society to correct it and increase linguistic awareness; whereas, according to the second principle, linguists should use data collected in a speech community for the benefit of the community itself. Wolfram's principle does not exclude Labovian principles but integrates them, adding a positive nuance to the researcher-community relationship, that the author defines collaborative (Wolfram, 1993:227). In Wolfram's idea, the ideal, but not exclusive, place where this principle can materialize is the educational context.

Motivated by such considerations, I initially take contact with the teachers of the classes (kindergarten and the first two years of primary school) who welcomed me and collaborators for data collection. I engaged in a dialogue with them about their needs and challenges they face in everyday teaching practices in superdiverse classes (see 6.2.1) and two main points emerged:

- the need for teachers for practical tools to manage the linguistic diversity of their classes.

- the necessity for children with a migratory background to discover spaces and occasions within the school environment allowing them to express their whole personality and identity, thereby revealing their domestic background.

These two points became the general guiding objectives of my proposal: presenting some tasks for managing linguistic diversity to the teachers; giving special attention to the "fund of knowledge and identity" of each student (Moll et al., 1992). Moreover, at least two corollary considerations, often stressed by teachers in our conversations, should be added: teachers face difficulties in integrating multilingual teaching strategies in regular school activities, mainly aimed at developing linguistic skills in Italian; bilingual children in their classes go through long phases of silence and states of inhibition in speaking their family language, during the early years of school (which De Houwer, 2021 describes as symptoms of non-harmonious bilingualism).

In search of activities that could meet these needs, storytelling seemed to be a suitable, motivating, and engaging option. Therefore, a workshop of 5 sessions focused on multilingual storytelling was designed and directed to 8 classes (2 in kindergarten, 6 in primary school) from two schools in the city of Bologna: I.C. 5 and I.C. 11. Moreover, a training course for teachers was also proposed, on the topics of linguistic diversity, multilingualism, and various teaching strategies to ensure bilingual children achieve harmonious bilingualism. Both the workshop and the course were conducted in spring 2022.

The workshop is presented in this chapter, which is structured into two main sections: section 2 explains the reasons for choosing multilingual storytelling as a working tool for the enhancement of linguistic diversity; section 3 presents the activities carried out in the five sessions. The chapter concludes with some final considerations on the impact of the workshop on the classrooms.

2. Why multilingual narratives?

The rationale based on existing literature for choosing a multilingual storytelling workshop is provided below.

2.1 Why "multilingual"?

The substantial international migration to Southern European countries in recent decades has brought significant social transformations due to its extensive scale and rapid growth. Until the 1970s, several of these countries predominantly witnessed great emigration to other European nations or continents. The abrupt reversal of the migratory pattern, coupled with the ethnically, religiously, and linguistically 'super-diverse' characteristics of immigrant communities, contributes to these changes (Bagna et al. 2007; Caruana & Scaglione, 2014). Multilingualism in our country has gained a new dimension through the phenomenon of international migrations. Caruana and Scaglione (2014) underscored that everyday contacts with multilingualism and linguistic diversity, resulting from immigration, are part of students' experiences, and it is within the school environment that students need to be guided in appreciating the value and potential of these linguistic resources. Otherwise, they argued that there is a risk that these individual and collective language skills could be perceived as largely external to the school community or viewed as challenges to overcome, while only European languages of broader communication receive recognition and prestige in the educational context. In addition, they found that often, bilingual and/or multilingual programs in South European schools are synonymous with the study of English, neglecting other languages that may be prevalent in the social context of Southern European countries.

Multi- and plurilingualism, i.e., the coexistence of multiple linguistic systems within communities (multilingualism) and individuals (plurilingualism)[73], is traditionally present in the

[73] For a discussion on this terminology see Luise (2013) and Galante et alii (2022).

Italian territory, thanks to numerous regional languages, varieties, and historical linguistic minorities (Fiorentini et al., 2020). Since 1975, Italian educational linguistics has developed reflections on democratic language education based on exploiting the student's entire repertoires. Heirs of the pedagogical reflections of Don Milani (Borg et al., 2013), in a country very different from the present-day Italy, those linguists and educators realized that the monolingual model for mass education was a symptom of social injustice[74] (Andorno & Sordella, 2017). The linguistic heritage carried by the children of the working class (e.g., non-standard varieties and regional languages) was neglected and poorly regarded in the school context, hindering the serene development of individual linguistic skills (Andorno & Sordella, 2017). The document "*Dieci tesi per un'educazione linguistica democratica*" (Ten theses for a democratic linguistic education), written by the GISCEL (Group of intervention and study in the domain of Educational Linguistics, 1975), guided by the linguist Tullio De Mauro, was the result of this line of considerations (Carbonara & Scibetta, 2020). It underlined how detrimental it is to neglect the individual's linguistic complexity and emphasized that effective educational interventions must involve the appreciation of the personal, family, and environmental background from both linguistic and cultural perspectives (Fiorentini et al., 2020). These considerations, born in reference to the gap between social classes, are still valid today, to guide educational practices in a linguistically superdiverse Italy, where languages which have a historical presence in the country share the same space with new immigrant languages (Bagna et al., 2007).

The achievement of full proficiency in Italian is a necessary condition for the success in school, for full participation and self-determination in society, of both recently immigrated students and students born in Italy from immigrated families. However, it is essential that acquiring adequate proficiency in Italian does not come at the expense of maintaining the language (or languages) of the families (Fiorentini et al., 2020). Research has demonstrated that multilingualism brings benefits both at the level of self-perception and of cognitive skills (see Cummins, 2000 on positive effects of

[74] See Garcìa et alii (2016:545-552) on the close relationship between linguistics (especially sociolinguistics), linguistic diversity, and social justice.

plurilingual education). However, cognitive positive effects emerge only when the entire repertoire is activated (Cummins, 2005) and extends beyond basic communicative skills (Rostald & MacSwan, 2018). It is crucial, therefore, that family languages do not remain "submerged languages" (Pugliese, 2015), confined to the realm of domestic communication, but become tools for cognitively challenging tasks, such as academic ones (Andorno & Sordella, 2021). The languages in the individual repertoire of bilingual individuals support each other, especially in the acquisition of literacy prerequisites. Building on Cummins' interdependence hypothesis (1979), numerous studies have shown that proficiency in the family language predicts and supports the acquisition of the second language, especially in sequential bilinguals (Gottardo et al., 2001 among others).

Bilingual children need to develop two languages and if they do so, their well-being is enhanced. De Houwer (2020) proposed that when subjective well-being is not negatively affected by factors relating to a bilingual setting, it's possible to speak of Harmonious Bilingualism. The counterpart of Harmonious Bilingualism is Conflictive Bilingualism. De Houwer (2020) also demonstrated that the well-being of young children is jeopardized when their home language is disregarded in early care, and they lack support in learning the societal language. (Pre-) schools play a vital role in fostering Harmonious Bilingualism by valuing children's home languages, enhancing their pride in these languages while simultaneously acquiring the societal language. When children perceive that their home language is disrespected in (pre-)school, they are more likely to reject it, feel unhappy in relation to their linguistic experiences and may show signs of depression, and be silent (De Hower, 2021).

Plurilingual education facilitate the process of enhancing each student's "fund of knowledge and identity" (Moll et al., 1992) in school (Payant & Galante, 2022). This expression refers to the set of knowledge and skills that parents and other members of the same community transmit to their children, which are not necessarily related to school literacy but allow individuals to function in a specific culture. This background is acquired during daily experiences that are historically and culturally unique and contribute to the process of defining identity. Language, its uses, and what is

associated with it can be considered part of this fund. A plurilingual critical approach assumes that all families, regardless of their origin, language, and even their socio-economic status, have competencies and knowledge produced through participation in various social practices, that can be transferred to their children (González et al., 2005). Those funds can be brought and positively exploited in classroom (Cummins, 2017).

This pedagogical approach places the learner's experience at the center, and it is advocated by various documents from European and Italian institutions[75]. Among these, of particular importance for the proposal presented in this chapter is the FREPA (Framework of Reference for Pluralistic Approaches; Candelier et al., 2011). It was developed by the Council of Europe to guide language education emphasizing pluralistic approaches, acknowledging, and utilizing the linguistic and cultural diversity of learners. Key characteristics include recognizing and valuing learners' multiple linguistic repertoires, promoting intercultural competence, and providing a framework for language teaching that goes beyond traditional monolingual approaches. It aims to foster inclusive language education that considers the rich linguistic and cultural backgrounds of learners. It offers descriptors to promote plurilingual and pluricultural competencies, which involve a person's ability to use multiple languages for communication and participation in intercultural interactions (Candelier et al., 2011).

Pedagogical strategies to guide students to achieve these skills are varied (see Andorno & Sordella, 2018, 2020 for educational proposals in the spirit of FREPA).

In "Plurilingual guide" developed by the Plurilingual Lab of McGill University (Galante et al., 2022), the authors present 5 plurilingual strategies (among many) to engage students' repertoire, that have been taken into consideration in the workshop presented below:

- **Cross-linguistic comparison**: comparing students' languages to other languages is an effective method to enhance engagement. This comparison can focus on linguistic features (phonology, syntax, vocabulary etc.) or language use.

[75] See Beacco et alii (2010) and Luise (2013) on plurilingual education in Europe and Saccardo (2016) on Italian legislative provisions on language education.

- **Cross-cultural comparison**: learning a new language opens doors to diverse cultures, emphasizing variations in customs, values, and language use. Cross-cultural comparisons, enhance critical thinking and broaden perspectives on knowledge.
- **Translanguaging**: translanguaging, the fluid use of different languages, is a potent method for comprehending and creating content in a new language. Learners can engage with diverse language materials, such as reading, writing, or listening, and then discuss and share the acquired knowledge in Italian or another language. This strategy promotes a dynamic and inclusive approach to language learning (García & Li Wei, 2014; García & Kleyn, 2016).
- **Translation for mediation**: whenever a new phrase, vocabulary, or grammatical element is introduced in the lesson, instructors can strategically plan activities that engage students in using the languages of their repertoire.
- **Pluriliteracies**: communication, seen as purposeful social engagement, is central to plurilingual instruction, acknowledging language learners as social agents performing various daily tasks with linguistic and cultural knowledge. In addition to traditional language skills, students employ diverse literacies, including visual elements, gestures, and digital skills. Instructors can promote versatile practices in any language, fostering both efficiency and meaningful understanding of new content and concepts (Cummins & Early, 2011).

These strategies allow for a change in perspective in everyday educational approaches. These are strategies through which all the knowledge in the classroom is activated, and new knowledge is built together. They are useful in maintaining a positive classroom atmosphere by valuing all aspects of each individual (Carbonara & Scibetta, 2020; Piccardo et al., 2021).

2.2 Why “narratives”?

The advantages of storytelling as an educational tool are numerous and can be condensed at least into three points:

- Positive effects on the overall cognitive and emotional development of the child
- Positive effects on the linguistic development of the child
- Positive effects on academic outcomes

Children naturally use stories as a platform to express thoughts, experiencing joy, reflection, elaborating sorrow, and change (see for example Caso, 2017 for a discussion on narratives with hospitalized children). Stories play a role in our lives from childhood to adulthood, seen in our engagement with entertainment like films, shows, and books (Kerry-Moran & Aerila, 2019). Scholars argue that children grasp concepts through stories even before understanding other logical structures (Bruner 1990), and that they think and learn through narrative forms (Bruner 1986, 1990). Life itself unfolds as a series of stories, blurring the distinction between whether our lives shape the stories or stories shape our lives (Bruner, 2002). Narratives are pervasive in children’s lives from an early age, and narrative competence is crucial to communicate social messages to others and to derive meaning from experience (Bruner, 1986; Nelson, 2007). It is also crucial in conceptual development (Stadler & Cuming Ward, 2005). Applebee (1978) highlights the role of stories in understanding temporal relationships, cause-effect connections, and theory of mind. Westby (1991) observes that narratives facilitate the use of language for monitoring, reflecting on experiences, and reasoning about, planning, and predicting future experiences. In addition, stories help socialization (Miller & Moore, 1989) through emotional attachment (Alexander et al., 2001). Story attachment is social because children create relationships with the characters in their favorite stories and because this emotional attachment emerge within the relationship with caregivers (Alexander et al., 2001).

The forms of narrative can be diverse, with the most studied in psycholinguistics being script, personal narration, and fictional stories (Bonifacio & Hvastja Stefani, 2010). The script represents the basic structure of every narrative, involving recounting a sequence of routine events linked chronologically or causally. It is the first form of narration that a child learns, through interactions with caregivers. Through the breakdown of these highly routinized, conventional, and familiar patterns, children learn to extract properties, object functions, and relationships between events (Nelson, 1985). With personal narratives, children learn to reflect on themselves and choose how to present themselves to others (Bruner, 1990; Puroila, 2019). Through fantasy narratives, they practice decentering and distancing themselves from their personal experience (Allen, 2012; Tartar, 2009;).

Narratives also serve as a crucial tool for oral language development, aiding in the use of more complex language than daily conversations require. Research supports the effectiveness of narratives in facilitating oral language skills, demanding explicit vocabulary, precise usage of pronouns, as well as of other referential expressions and mastery of temporal connectives (Stadler & Cuming Ward, 2005). Active engagement in storytelling activities helps develop the child's pragmatic skills, specifically referential abilities. Narratives provide opportunities for children to listen to and use linguistic means, including articles and other referential expressions, in the construction of complex texts that are also motivating (see 3.1.4). Implementing educational practices centered around storytelling in the classroom allows children to draw on all available linguistic resources in order to be communicatively effective. Storytelling activities fall within those linguistically relevant practices that appear to be beneficial to the child's morphosyntactic and pragmatic development (see 5.3).

Models of narrative development are many (Applebee, 1978; Karmiloff-Smith, 1979; Paul et al., 1996; Stadler & Ward, 2005; Westby 1984; among others), but briefly described, the physiological development of narrative skills stretches from a pre-narrative stage (around 2 years), where the child labels objects or events, to a narrative stage that resembles adult-like narratives around 9 or 10 years. The intermediate phases of interest in this chapter are those between 4 and 7 years, the age range of the children in the classes I worked in. At 3-4 years, the child introduces new elements into the story,

and the first fantasy narratives appear. At 4 years old, the child can retell a story heard before, respecting the plot if images are present, but may not yet master the narrative structure. Between 5-6 years old, stories become richer, with sequences connected by causal, temporal, and referential relationships, although not always appropriately. After 6 years old, the basic narrative structure seems acquired, and complexity increases with linguistic competence (Berman & Slobin, 1994; Bohnacker & Gagarina, 2020; Gagliardi, 2019).

In preschool years, narrative competences are considered to be an important predictor of later development of literacy (Wellman et al., 2011). Specifically, narrative skills are thought to be related to reading comprehension (Bonifacci et al., 2018). Proficiency in storytelling therefore is crucial for academic achievement, as it serves as a means for sharing and acquiring academic information, connecting oral language to literacy (for a recent collection of studies on the topic, see Veneziano & Nicolopoulou, 2019). These skills are considered essential, bridging the gap between spoken language and literacy by exposing individuals to extensive, contextualized, cohesive discourse units and abstract texts. Narrative skills expose children to longer and structurally more complex texts, providing an excellent source of rich linguistic input. Moreover, third-person narrations place the child in front of stories that are distant in time, space, and from the child's own experience, serving as exercises in distancing and decentering. All these skills are necessary to access disciplinary texts and for subsequent academic success (Gagarina et al., 2016).

Distancing and decentering are said to be important abilities also for the acquisition of intercultural competences (Candelier, 2011:23). Stories can expand children's views so that they consider moral and socio-emotional issues beyond their personal experiences and ponder from a safe distance what choices they would make in similar circumstances (Karry-Moran & Aerila, 2019:5). Stories are tools empowering children and enhancing their moral development through increasing cultural awareness and sensitivity. Using stories to teach compassion and develop empathy is vital from the perspective of coping in an ever more diverse world (Pesonen, 2019).

There is a strong connection between narrative education and multilingual pedagogy. Cummins (2005) suggested the creation of bilingual stories in the classroom as an effective strategy for recognizing students' family languages as educational resources.

Carbonara et alii (2023) investigated whether exposure to multilingual pedagogies enhances emergent bilingual children's narrative abilities, by comparing two groups of emergent bilingual children with migrant background attending public primary school in Italy. 30 children were exposed to multilingual pedagogies, whereas 33 attended a traditional monolingual program. The findings indicate that children who experience multilingual pedagogies incorporate a higher number of mental state terms in their narratives, both in their family language and Italian, compared to those exposed to monolingual education. Their findings also reveal a dynamic interaction between the home and school languages in the narrative construction among children exposed to multilingual pedagogies. This study highlights that multilingual pedagogies enhance the literacy skills of emergent bilinguals, allowing them to draw upon these skills across all languages in their repertoire.

The decree by Italian Ministry of Education, *Indicazioni Nazionali per il curricolo della scuola dell'Infanzia* (d.m. 16 novembre 2012, n. 256) also places particular emphasis on narration and reading stories from different cultures, when listing various methods for recognizing linguistic diversity and enhancing bilingual children well-being. Narrative content and structure are greatly influenced by culture (Carmiol & Sparks; 2014; Wang & Leichtman, 2003) and that makes stories a privileged tool for promoting multilingualism and intercultural dialogue.

The importance and effectiveness of integrating storytelling as a regular component in the preschool curriculum are demonstrated by various studies, particularly for children from low-income families (Nicolopoulou et al., 2015; Nicolopoulou, 2019; Flynn, 2019).

3. Multilingual narratives workshop

The workshop consists of 5 tasks, which were conducted in 5 sessions during regular class hours. Each session was tailored to the school's hour unit (typically 45 minutes). The task execution didn't always span the entire hour, considering the young age of the children (4-7 y.o.) and their attentional resources.

Before implementing the teaching proposal, each teacher was asked to provide a list of all the languages present in the repertoires of the children. The languages involved in the project are (as reported by teachers): Albanian, Arabic (North African varieties), Bengali, English from Ghana, and Nigeria, Ewe, Filipino, French (African varieties), Mandarin Chinese, Romanian, Russian, Sinhalese, Spanish (South American varieties), Tigrinya, Ukrainian, Urdu, and Italian.

The workshop was designed with an increasing complexity of narrative tasks:

- Task 1: Lexical ice-breaking and labeling task
- Task 2: Personal storytelling
- Task 3: Listening to and understanding a fictional story
- Task 4: Retelling a fictional story
- Task 5: Telling a fictional story

Each task includes a description of plurilingual strategies, approximate duration, goals related to narrative competence, goals related to plurilingual competence (as detailed in FREPA), materials needed, instructions and some final considerations about the activity.

A caveat is necessary in this regard: this workshop is meant as a guide and a collection of suggestions for teachers. Multilingual activities should be repeated throughout the year and integrated into regular school activities, rather than limited to isolated workshop experience. Each teacher could utilize the proposals as a template for creating their own tasks, beyond the workshop. Furthermore,

to implement an inclusive critical plurilingual approach, every activity should be adapted to the specific school context.

TASK 1: Lexical ice-breaking and labeling task

Plurilingual strategy	Translanguaging; Translation for mediation; Pluriliteracies
Approximate duration	30 min
Goals	familiarizing with researchers; labelling entities in the languages of the class repertoire; creating a conducive environment for the use of family languages in the classroom; engaging members of the communities to which the families of the children belong in school activities
Material	Ball
FREPA Competences	*K 2.2* Knows that each individual belongs to at least one linguistic community and that many persons belong to more than one linguistic community *K 5.1* Knows that there are very many languages in the world

The task consists of multiple steps, each involving a simple activity or game. Its fundamental purpose is the gradual familiarization of children with unfamiliar individuals present in the classroom and the lowering of emotional filters in the use of all languages of the class repertoire. In the classroom in addition to children and the teacher, the researcher was also present and together with some language mediators, who speak Mandarin Chinese, Spanish, Moroccan and Egyptian Arabic, all from the University of Bologna. The room needs to be adapted for activities that require an empty space (teachers could move to the gym or move the desks to the wall).

Step 1. Around the World in Eighty Greetings: everyone will form a circle. The activity-conductor holds a ball (which, in this case, was also a world map), says his name, and accompanies it with a greeting in a language from his personal repertoire. Then, the conductor passes the ball to the person on his right. The person holding the ball speaks and does the same. The activity continues until it has been carried out by all the people in the circle.

Step 2. Around the World in a Thousand Greetings 2: The round of names and greetings is repeated, but using a greeting proposed in step1 by another person in the circle.

Step 3. The Greeting Airplane: While remaining in a circle, everyone stands up. All participants open their arms as if they were airplanes. All together they decide on the destination of the trip, which must be a country where one of the languages of the class is spoken. After mimicking the airplane journey (fastening seatbelts, takeoff, etc.), they simulate landing and getting off the plane. Once they "arrive" at their destination, children walk around the space, "visiting the new country", and when they meet, they must greet each other in the language of the country. If more than one language is spoken in the country, multiple greetings can be used. Before takeoff, the airplane commander (the game director) repeats which languages are spoken in the country and what greetings are used there, with the help of the children.

Step 4. Arriva un camion carico, carico di...: This is a classic Italian children's game to learn names of animals, which has been revisited in a plurilingual version. The activity-conductor starts by saying "arriva un camion carico carico di …" (here comes a truck loaded with ...) and adds the name of an animal in Italian, to which the children must respond by imitating the sound or making a movement that resembles the mentioned animal. Then, the language mediators follow the same pattern, but at the end of the formula, they mention the animal in their mother tongue. At this point, only the children who belong to the same linguistic community will understand which animal it is, and they will be the first to propose the sound/movement, serving as models for other classmates to follow. After a few examples, when it is made sure that the children have understood how the game works, they should lead the activity themselves and propose animals in the languages of their repertoire, including Italian.

Step 5. Touch the ...: The game follows the same pattern as the previous one, but animals are replaced with body parts, and the opening formula is "touch the ...". Each participant must respond to the formula by touching the mentioned body part.

The activities have shown several potentials. First of all, they are fun games that help lower emotional filters and create a positive atmosphere in the classroom, which is particularly important

with kindergarten and early primary school children and those going through a silent phase. In this type of population, linguistic performance in structured activities is highly dependent on emotional and environmental factors (Daloiso, 2015). Furthermore, the structure of the activities is simple, easy to understand, and not demanding in terms of linguistic production; it only involves producing single words, highly ritualized like greetings or very familiar terms, such as those designating animals and body parts. All the games require a gradual involvement of the child and a bodily response as well. Linguistic production is free, never forced (contrary to question-answer schema) and when a verbal answer is required, Italian is always accepted. Steps 4 and 5 focus mainly on comprehension, and production is left to the child's free initiative. The presence of a member of the linguistic communities of some children has also proven to be a decisive factor for their involvement and motivation and has ensured a smooth transition from a monolingual Italian-centric environment to a multilingual environment where children feel entitled to use other languages. If it is not possible to involve community members, teachers can use online tools to find translations of vocabulary into the children's languages and use online voice synthesizers to make the word heard during the game. During the classroom activity, a list of prompts in languages for which a native speaker was not available was also used and it was previously prepared and then spoken by the adults. The risk is that incorrect pronunciation by the non-native adult may compromise the child’s comprehension. Involving native speakers is therefore recommended. The presence of members of the speech community will be a constant in all tasks of the five sessions.

Vocabulary (greetings, animals, and body parts) was chosen within the semantic classes that constitute the basic lexicon learned early by children, even in a domestic context, and thus in family languages. Moreover, the chosen words pertain to very familiar semantic domains, often present in children's stories cross-culturally. It was also possible to use materials present in the classroom, such as plastic toy animals and models of the human body. In this way, the linguistic experience is anchored not only to the body (through the stimulation of bodily responses) but also to real objects with which the child has a positive and daily relationship.

Favaro et alii (2012) recommend setting aside a weekly time for teaching to use words to name, describe, respond, take the floor. Ideally, having three days a week for a couple of hours each time would allow dividing the children into small groups and carrying out lexical enhancement activities, such as the one presented here. Lexical development is crucial for the child's overall linguistic development, including grammatical development (Bates & Goodman, 1999), and it is in line with such observations that this first activity was proposed.

TASK 2: Personal storytelling

Plurilingual strategy	Translanguaging; Translation for mediation; Pluriliteracies
Approximate duration	30 min
Goals	Narrating personal and repeated events; organizing events of one own's daily routine chronologically; using words and phrases from the domestic and personal domain in the family language at school.
Material	none
FREPA Competences	*K 2.2* Knows that each individual belongs to at least one linguistic community and that many persons belong to more than one linguistic community *K 5.1* Knows that there are very many languages in the world

The session (like all the following ones) opens with the greetings learned in the first session, each child choosing the one they prefer. The second task involves two steps:

Step 1. Let's "build" our morning: the activity can be conducted together or in small groups. Children and adults form a circle. The purpose of the activity is to tell about one own's morning, from waking up to entering school. Each group member must mention an activity accompanied by a gesture, while others repeat the expression and gesture. Each action can be mentioned in a different language. The activity begins with the game-conductor proposing the first action in Italian, followed by mediators with actions in other languages, and then letting the children continue. Each participant must mention an activity chronologically following the one just mentioned.

Step 2. Watch and repeat: the conductor, one of the mediators, or the children (in turns) repeats a daily action using the expressions presented in the previous phase; others must respond with the corresponding gesture. Participants are spread throughout the classroom, and those who make a mistake (responding with the wrong movement) sit down until only one winner remains. For this second phase, it is recommended to select a set of a few expressions from those mentioned in step 1.

Considerations made for task 1 regarding the presence of native speakers and the preparation of prompts in the various languages of the students remain valid for this activity as well.

This activity requires the collaborative construction of a script. As shown in section 6.2.2, it is one of the first type of storytelling which children are exposed to, and it is less challenging for them (Berman, 2009). Through routinized actions that mark their day, children come to build in their minds the concept of chronological time, to which a wide range of linguistic expressions is recurrently associated (Daloiso 2009: 102). The day, marked by a fairly predictable use of time and space and the repetitiveness of actions, constructs a rather transparent and safe framework within which all children, even those who are not yet full competent Italian speakers, can place themselves without getting lost and disoriented (Favaro, 2012). For this reason, it is advisable to start a narrative-focused classwork with stories anchored to routines and that do not require distancing from personal experience. Moreover, morning actions take place at home, the privileged place for the use of family languages, and therefore constitute verbal vocabulary that the child is likely to know in his family language. It is an excellent bridge to bring the experience and knowledge accumulated at home to school.

It should be emphasized that in the transition from kindergarten to primary school, the change in routine is one of the most significant changes in a child's life and of his family and can be a source of stress (Ferretti & Bub, 2016; Wildenger et al., 2008). Activities that focus on the structure of the morning routine help children better understand and internalize new, still unfamiliar patterns. The

process of verbalization contributes to reinforcing the connection between physical action and its conceptual meaning, thus facilitating the learning and memorization of new knowledge.

TASK 3: Listening to a fictional story

Plurilingual strategy	Cross-linguistic comparison, cross-cultural comparison
Approximate duration	40 min
Goals	Listening to and understanding a fictional third person story; individuating the main elements of a narrative: characters, time, and space; organizing the episodes of a fictional story chronologically;
Material	Internet connection, interactive whiteboard, work sheet with the episodes to be ordered.
FREPA Competences	*K.2.2; K 5.1* *K 5.3* Knows that there are many different kinds of script

This task involves three steps:

Step 1. Listening: students listen to a story with the support of pictures projected on the interactive whiteboard.

Step 2. The three corners: children are divided into three small groups, each of which will work on a different short activity. Each activity is assigned a space in the classroom, and at the end of each, the group moves on to the next activity, rotating through different corners of the class:

- Corner 1: the group is provided with pictures from the story just heard. Guided by the teacher/instructor, they must orally identify the protagonists (who), the environmental setting (where), and the temporal placement (when) using prompting questions like: e.g., who is in the story/who are we talking about? where is X? When does X happen? The answer to "when" is usually found in the typical opening formulas of stories (e.g., once upon a time...) that place events in an unspecified time. The conversation can take place in all languages of the repertoire.
- Corner 2: chronological ordering of the story sequences (pictures provided).

- Corner 3: each group member receives a sheet with a printed flower, on the petals of which a word (related to the central theme of the just listened story) is written in different languages. Each word is dotted, and the children are asked to trace it (see Appendix D).

The stories heard in class were selected from two websites[76] that collect stories from around the world, filterable by age, difficulty, and language. These resources present the same stories in different languages, both oral and written. We started by playing the audio in Italian and tried to structure the activity dialogically, stimulating children's interventions and leaving room for their questions, following the dialogic reading model of Lonigan and Whitehurst (1998; see also Towson et al., 2017).

Before the start of small group activities ("corners"), time was allocated to explore the digital resource together with the children. They were interested in finding their family language in the list of selectable languages in the webpage and seeing how they appear on the screen. In some classes, the story was listened to again in the requested languages, while in others, the various writing systems were observed, focusing on two points: comparison with the graphemes of the Latin alphabet, with which the children are becoming increasingly familiar, and the direction of writing. This moment proved to be very valuable because it allowed everyone to make a personal contribution. The children were able to recognize writings in their own language. Interesting discussions arose about different languages using the same alphabet and that some languages are written in directions different from Italian.

In one class (first grade), this led to questions about how books written in similar alphabets are handled and read, and the answers came from the students themselves. We observed how books in the Latin script are structured, focusing on their components such as the cover, introductory pages, and especially the orientation of writing and reading. Some children were able to compare these observations to their personal experience of handling books in the Arabic alphabet and were proud o share such considerations with their classmates. The resources used were therefore particularly

[76] Storybook Canada: https://www.storybookscanada.ca; Storyweaver: https://storyweaver.org.in/en/

suitable for developing awareness on main conventions of print[77], in a multilingual dimension (Witherhurst & Lonigan, 2001; Akoğlu & Kizilöz, 2019).

Teachers showed enthusiasm when even usually inhibited and shy children spoke up, wanted to write, or point to their language script. Particularly interesting was the case of children with Pakistani background who were learning to read in Arabic because they attended the local Quranic school. They wanted to approach the whiteboard with the text of the story projected in Arabic, pointing to the letters, and saying letters names aloud, even attempting to read some syllables, in front of the whole class. This is exactly what is meant by activating personal knowledge in the classroom (see Andorno & Sordella, 2017:200 for similar experiences).

This adds to the fact that the proposed stories come from all over the world, and therefore, the task was an opportunity to show teachers that multilingual narratives can lead to interesting discussions that go beyond linguistic forms and focus on content, also in a comparative perspective.

This activity also served as an introduction to third-person stories, which are particularly challenging for children. Achieving full competence in this type of narrative takes time and spans the entire primary school cycle. This type of stories requires mastery of various linguistic devices, including cohesive means and referential expressions. Furthermore, it presupposes the ability to distance oneself from one's personal experience. While the activities of corners 1 and 2 were in line with the previous tasks, the activity of corner 3 served to connect the spontaneous considerations that emerged during the observation of different scripts to a practical activity. Additionally, for children who are learning to write, practicing fine motor skills by tracing letters (although in an alphabet different from the Latin one) is an excellent exercise and is essential for gaining confidence in writing practice (Chandler et al., 2021; see also Tolchinsky, 2008; Ball, 2008).

[77] According to Whitehurst and Lonigan (2001) books adhere to a set of conventions that can be comprehended even in the absence of reading skills. In Italian, these conventions encompass the left-to-right and top-to-bottom orientation of text on each page, the progression of print across pages from front to back, the distinction between book covers and pages, the differentiation between images and text on a page, and the interpretation of punctuation elements, including word spaces and sentence-ending periods. Familiarity with these conventions contributes to the reading acquisition process. This set of knowledge is part of what the authors call print awareness, which is one of the main components of emergent literacy (Lonigan and Whitehurst, 1998:850).

TASK 4: Retelling a fictional story

Plurilingual strategy	Translanguaging; Translation for mediation; Cross-linguistic comparison
Approximate duration	40 min
Goals	Developing children's printing awareness; identifying the main components of Story Grammar; retelling in Italian an episode heard in the family language.
Material	Internet connection, interactive whiteboard, work sheets
FREPA Competences	*K.2.2; K 5.1* *K6.5* Knows that each language has its own phonetic / phonological system

This task is based on the story grammar elements identified by Stein and Glenn (1979), according to which a good story should contain: setting (introduction of the protagonist and the physical context in which the events take place); initial event (that generates a problematic situation); internal response of the protagonist (psychological state that motivates the protagonist to devise a strategy); attempts (implementation of the strategy); consequence (positive or negative result of the protagonist's actions); reaction (feelings of the protagonist related to the consequence); ending (resolution of the problem). The story chosen for this task is titled "*La pecorella vanitosa*" (The Vain Little Sheep), which is an unpublished story constructed by a class from the Don Bosco primary school (Polverigi, Ancona, Italy).

The first phase of the task involved the preparation of materials. The narrative consists of 9 hand-drawn pictures, scanned, and arranged in sequence on a PowerPoint presentation. The text consists of 9 sequences (see Appendix D), each translated into a different language with the help of native speakers (university students, parents of the children, shopkeepers in the neighborhood of the schools) (see Appendix D). Each native speaker was then asked to record an audio while reading the sequence in their language and send it to me. Audio in a different language that narrates the episode was added to each slide of the PowerPoint. Some episodes were recorded in multiple languages, and various PowerPoints with different language combinations were prepared, adapting to the linguistic composition of the different classes. The task included the following steps:

- **Step 1**: observe the "cover" (first slide) and identify its constituent elements: author and title[78] (orally conducted)

- **Step 2**: develop expectations about the story based on the title: what will the story be about? (orally conducted)

- **Step 3**: observe the first picture of the story and listen to the first sequence (always in Italian).

- **Step 4**: look at the following pictures, listen to the story, and guess the languages.

- **Step 5**: translate the content for the whole class by those who understood it (sequence by sequence)

- **Step 6**: verify comprehension through coloring the pictures (worksheet provided; see Appendix).

- **Step 7**: (only in primary school) rearrange the episodes chronologically and match each sequence with the element of the story grammar it represents (worksheet provided, see Appendix D).

This task proved to be highly motivating and engaging for all children and was particularly suitable for making everyone feel valued. In fact, after listening to a sequence in a language, the child who speaks that language would activate and become an indispensable resource for the understanding of the entire narrative by the whole class, being the only one able to translate the episode. The story, in this regard, has several advantages. Being based on a repetitive structure, given the identical narrative patterns, it was easily understandable and predictable. The images also provided significant support for understanding. Even the children with lower proficiency in their family language and who did not understand well could still tell the story to the class, inferring what happened from the always-similar structure or from the images. Similarly, the child with limited proficiency in Italian could be linguistically scaffolded by the peers, who could understand the story by observing the pictures. It is also particularly suitable for working on the story grammar of Stein and Glenn (1979) because it reproduces it straightforwardly.

[78] These are part of print conventions.

The vocabulary that recurs in the story is about animals, household objects, and colors. It can also be a good tool for working on lexical interlinguistic comparison, even in foreign language classes (English in the case of Italian schools).

The first steps of the activity aimed to work on print awareness, focusing on aspects external to the narration, such as author and title, and then asked the children to develop expectations about the story based on the title. Developing expectations before facing the story improves comprehension and proves to be a good habit to acquire during literacy learning (Ammon, 1975).

The activity also aimed to let children hear how different languages sound, creating an opportunity to work on crosslinguistic awareness and understanding that languages are also different in the phonetic dimension. It is interesting to note that even reproducing the scene in a language not present in the class repertoire often led to moments of interesting metalinguistic analysis by the children. For example, an Italo-Ukrainian girl, when asked "what language is the one you just heard?" after listening to an audio in Albanian, answered “English”. She justified her answer based on the "strange r sound" that she heard during the story. It should be noted that, indeed, the alveolar approximant [ɹ] is common to the phonetic inventories of both English and Albanian.

Being a story based on colors, which become the central element of each episode, it is possible to verify comprehension by asking the children to color the images appropriately, which were printed in black and white, creating a specific sheet. The sheet then required to arrange the images in sequence and match them with the corresponding label: beginning, attempt 1, consequence 1, attempt 2, consequence 2, attempt and consequence 3, attempt 4, consequence 4, end. Teachers were then tasked with reusing the proposed story grammar for the analysis of other stories during their regular teaching.

The activity is structured in a cooperative way, and the child who knows the language feels like an essential resource for achieving the collective goal. Without the contribution of that specific classmate, the group could not have achieved the goal of the activity, i.e., the understanding and full enjoyment of the story. This creates a space for cultural and linguistic valorization through individual empowering. Languages are not disembodied, and valorization passes through personal experience.

Children, whose language is reproduced, are indirectly encouraged to speak; in other words, their active participation is sparked by personal initiative (not forced), driven by the connection between the child and the language. Their languages are brought into the classroom and heard by everyone; the classroom becomes a place for sharing personal experiences and valorizing individual knowledge.

TASK 5: Telling a fictional story

Plurilingual strategy	Translanguaging; Translation for mediation
Approximate duration	30 min
Goals	Telling a well-formed story using the entire linguistic repertoire of the class.
Material	Internet connection, storytelling dice
FREPA Competences	*K.2.2; K 5.1; K 5.3; K 6.5*

After providing various story models, testing productive skills through retelling a simple episode, and collectively examining the grammar of the story, the final activity saw the children become storytellers. The task can be conduct in small groups or with the whole class. It involved the following steps:

Step 1. Choosing the initial temporal formula together.

Step 2. Three story-telling dice were given- one with a different setting on each face, one with different protagonists, and one with different villains or enchanted objects (see Appendix D). The children are asked to roll the dice and invent a story following the story structure introduced in the previous task and using elements determined by the dice roll. Each child is then encouraged to contribute to the narration not only in Italian but also in other languages (whether they are foreign languages studied at school, languages spoken at home, or learned in other contexts).

Step 3. Reading the story aloud in class.

For the first step, a list of opening formulas for fairy tales from different linguistic traditions was presented (see Appendix D). These were read together, supplemented by the children's suggestions, and chosen as the beginning of the story to be constructed together.

The activity was met with enthusiasm from both teachers and students. Primary school children demonstrated the ability to reflect on the basic structure of a well-formed narrative, even after just a few dedicated activities. Of particular importance is the supportive attitude that the children showed in constructing the narrative and their interest in hearing words from their classmates' languages. This supportive atmosphere led all children to contribute, both in Italian and in their family languages. The final reading aloud of the result, proved to be a moment of strong class cohesion (see Appendix D for an example of the final story created by children).

4. Final remarks

The workshop design, for feasibility reasons, did not include feedback collection, hence it is not possible to measure the achievement of the objectives. In particular, it is impossible to evaluate the effects of the activity on the specific linguistic skills related to storytelling. However, it is possible to gather some observations on the impact regarding the objective of enhancing multilingualism.

The activities were enthusiastically received, confirming that storytelling is an enjoyable and engaging activity for children. Children who were usually silent or inhibited to speak their family language actively participated in the tasks, even just producing isolated words in the family language, especially in the last phase of the workshop. This result appears to be linked to the positive classroom atmosphere. The supportive attitude of classmates and the interest in cultures and languages shown by everyone, including teachers, contributed to the well-being of children, and lowered emotional barriers.

Teachers were enthusiastic about the workshop and requested the material used. Therefore, a plurilingual education kit was prepared, collecting the sheets, digital resources, and readings used to

design and conduct the workshop (the material used for the workshop and shared with teachers is included in Appendix D).

The importance of involving members of the speech community has emerged clearly. They are carriers of specific languages and cultures and brought them into the school context, still traditionally monolingual (Chini & Andorno, 2018). School, in this way, becomes a catalyst for different forces, resources, and knowledge, working synergistically to educate children and ensure their well-being, both inside and outside the school (Henderson & Mapp, 2002). Storytelling can be an excellent way to bring the voices and experiences of linguistic minorities into the classroom, both for immigrant languages, as in the case presented here, and for historical Italian regional languages. Stories could serve as carriers of fading oral traditions, including those associated with Italian dialects, enabling individuals to traverse both space and time. They facilitate a journey of distancing ourselves and connecting not only with those diverse in origin but also in age.

Multilingual storytelling activities also seemed suitable for stimulating active participation of children. Kangas and Reunamo (2019) indicated that when children's ideas are listened to and their actions change the situation in pedagogical activities, they experience participation. Participation involves children realizing that their voices matter, and their knowledge can influence the direction of activities. By engaging in interactive storytelling, children's agency flourishes, and they actively contribute to collaborative knowledge-construction processes. A shy child who may hesitate to take the floor can learn to articulate their ideas and take initiative (Kangas & Reunamo, 2019; Payant & Galante, 2022).

The workshop seemed to confirm that the combination of plurilingual pedagogies and storytelling activities is a powerful and effective mean of inclusion and can lead to the emergence of dynamics promoting empowerment among students with immigrant background in super-diverse classes.

Conclusion

Stories seem to be pervasive in the lives of children and are the preferred means through which children explore external and internal worlds, learning to verbally articulate their experiences. Narrative productions thus serve as a suitable lens for examining children's linguistic development and as an effective pedagogical tool for supporting it, particularly when this development is characterized by the interaction between multiple languages and cultures. This is evident in the case of bilingual children born in Italy and coming from families who have immigrated from other countries, speaking a language different from Italian at home and transmitting diverse linguistic and cultural heritages.

These children were the focal point of the work presented in this thesis. The investigation aimed to contribute to understanding bilingual linguistic development and the pedagogical practices that can support it. Narrative was chosen as both a data collection tool and an educational instrument due to the significance of stories in children's lives and their familiarity with this text type.

The research focus was on the expression of definiteness, crucial for constructing a cohesive and easily navigable story, making it an intriguing subject for studying children's pragmatic skills. The adopted theoretical perspective views definiteness as a grammatical category aimed at encoding a category of meaning known as identifiability. Identifiability is a language-independent category to which children appear to develop early sensitivity, whether monolingual or bilingual. The formal expression of this category varies significantly from language to language, and bilingual children speaking languages with vastly different identifiability encoding systems may encounter challenges.

The study, in line with previous findings, highlighted a crucial interaction between pragmatic skills, and the morphosyntactic abilities of bilingual children. The bilingual children involved in this study came from family backgrounds differing in terms of use of languages at home, exposure to Italian, and involvement in linguistically relevant activities. From this perspective, Arabic and Spanish-speaking children appeared very similar, demonstrating comparable linguistic abilities. Even

in the expression of definiteness, the two groups behaved similarly and closely resembled monolingual peers, partly due to their greater dominance in Italian and the proximity of identifiability encoding systems between the family language and Italian. In contrast, the Chinese-speaking group differed significantly, both in terms of linguistic experience and general expressive abilities in Italian, showing a clear disadvantage compared to peers in the other bilingual groups. Chinese children were more dominant in the family language, and its strong influence on Italian in the specified domain was evident. Nonetheless, they showed sensitivity to the dimension of identifiability, although expressed with different formal means. Interestingly, the latter are influenced by functionally analogous means found in the family language, resulting in a greater mastery of indefinite over definite articles.

All children then benefited from the multilingual narrative activities proposed in the classes, both individually, as a moment of personal empowerment through the enhancing of their linguistic background, and collectively, offering an opportunity to explore diverse cultural specificities and linguistic complexity, leveraging it in the classroom. Teachers also benefitted from the proposed program, gaining knowledge of educational tools for managing linguistic and cultural diversity and updating their skills.

The study presented in the preceding pages has some limitations, primarily the small number of participants, limiting the statistical analysis's power and generalizability. Additionally, the pandemic conditions during the research project, imposed restrictions that affected data collection. Contact with families was indirect, and intermediate steps often led to a loss of information from the questionnaires. It was also not possible to administer tests to the children on more executive functions, which play a fundamental role in the development of the targeted linguistic abilities. Another significant limitation is the lack of balance in the groups based on the age of first exposure to Italian, attributed to relaxation of selection constraints, coming from the difficulty in recruiting actively bilingual children in the family language across three different linguistic communities.

Prompts for future research can also be identified. The group of Italian-Chinese children in Italy proved particularly interesting, both in terms of linguistic development and their unique

linguistic experiences. The preference for indefinite over definite articles, attributed in the study to the influence of the first language, would be interesting to explore further with a longitudinal study involving younger children, aiming to reveal the order of acquisition of Italian determiners in this specific population.

This population is unique, and it remains uncertain whether the typical dominance switch seen in heritage learners will occur (Montrul, 2016). Given the variability of language proficiency and the trajectories of language development among bilingual speakers, some learners may achieve near-native proficiency in both their languages, while others may struggle to maintain the family language. Italian-Chinese bilingual children still seem to be at a school-age close to L2 Italian speakers, despite early exposure. Thus, it is an interesting developmental trajectory to follow across the lifespan.
Specific studies of a more experimental nature on processing mechanisms, on the link between cognitive load and article omission, and on Cross-Linguistic Influence (CLI) would be desirable. Experimental studies are specifically encouraged for both Chinese and Arabic speakers, where the article domain appears as a good field to test hypotheses on CLI, as well as in other domains. Moreover, in a follow up study, it would be important to determine how native adult and child speakers exposed to the participants narrative would react. Would they detect omissions and also non-native usages of determiners?

Finally, the presented educational activities should be repeated, and more structured feedback from both children and teachers collection should be conducted, also using qualitative methods.

Abbreviations

ACC: accusative
ADV: adverb (-ial)
ART: article
CL: classifier
CRS: currently relevant state
DEF: definite
F.: feminine
GEN: genitive
INDEF: indefinite
M: masculine
NOM: nominative
O: object
PERF: perfect
PFV: perfective
S: subject
SG: singular
V: verb

References

Abbott B. (2010). *Reference*, Oxford University Press, Oxford.

Abi Aad, A. (2006). Insegnare l'italiano ad arabofoni, variazioni e interferenze [Teaching Italian to Arabic speakers, variations and interferences*]*. *Rassegna Italiana di Linguistica Applicata*, 323-343. 38, 2-3. Milano: Bulzoni.

Adams, A.-M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language & Communication Disorders, 35*(1), 95–116. https://doi.org/10.1080/136828200247278

Aguilar-Guevara A., Pozas Loyo J., Vázquez-Rojas Maldonado V. (2019), *Definiteness across languages* (Studies in Diversity Linguistics 25), Berlin: Language Science Press.

Akoğlu, G., & Kizilöz, Ç. (2019). Print Awareness Skills and Home Literacy Environment of Turkish Preschoolers. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 34(1), 90-105. https://Doi.Org/10.16986/Huje.2018040772

Aleksynska, M., Tritah, A. (2014). Occupation-education mismatch of immigrant workers in Europe: Context and policies. *Economics of Education Review*. 229-244. 36(C).

Alexander K.J., Miller P.J., Hengst J.A. (2001). Young Children's Emotional Attachments to Stories. *Social Development*. 375-398. 10 (3). https://doi.org/10.1111/1467-9507.00171

Allen, S., & Schröder, H. (2003). Preferred argument structure in early Inuktitut spontaneous speech data. In J. Du Bois, L. Kumpf, & J. Ashby (Eds), *Preferred argument structure: Grammar as architecture for function.* 301–338. Amsterdam: John Benjamins.

Allen, S. E. M., Skarabela, B., & Hughes, M. E. (2008). Using corpora to examine discourse effects in syntax. In H. Behrens (Ed), *Corpora in language acquisition research*, 99-137. Amsterdam: John Benjamins. DOI: 10.1075/tilar.6.07all

Allen, J. R., Allen, S. F., Latrobe, K. H., Brand, M., Pfefferbaum, B., Elledge, B., Burton, T., & Guffey, M. (2012). The power of story. *Children & Libraries: The Journal of the Association for Library Service to Children*, 10(1).

Ammon, R. (1975). Generating Expectancies to Enhance Comprehension. *The Reading Teacher*, 29 (3), 245-249. https://www.jstor.org/stable/20193996

Andorno, C. (2010). Lo sviluppo della morfosintassi in studenti cinesi [The development of morphosyntax in Chinese students]. In Rastelli S. (Ed). *Italiano di cinesi, italiano per cinesi: dalla prospettiva della didattica acquisizionale, 89-122. Guerra.*

Andorno, C. (2011). La grammatica per l'apprendente di L2. Apprendenti guidati e spontanei a confronto. [Grammar for the L2 learner. Guided and spontaneous learners compared]. In Corrà, L., Paschetto, W. (Eds*). Grammatica a scuola*, 36-48. Franco Angeli.

Andorno, C., Sordella, S. (2017). Esplorare le lingue in classe. Strumenti e risorse per un laboratorio di Eveil aux langues nella scuola primaria [Exploring languages in the classroom. Tools and resources for an Eveil aux langues workshop in primary school]. *Italiano LinguaDue,* 9(2), 1-68. https://hdl.handle.net/2318/1660086

Andorno, C., Sordella, S. (2018). Usare le lingue seconde nell'educazione linguistica: una sperimentazione nella scuola primaria nello spirito dell'Éveil aux langues [Using second languages in language education: an experiment in primary school in the spirit of Éveil aux langues]. In De Meo, A., & Rasulo, M. (Eds). Usare le lingue seconde, 211-233. AItla, Milano.

Andorno, C., Sordella, S. (2020). Noi e le nostre lingue. potenziare attitudini metalinguistiche in laboratori di éveil aux langues [Noi e le nostre lingue. Enhancing metalinguistic aptitudes in éveil aux langues workshops]. *Italiano LinguaDue,* 12(1), 330-352DOI: https://doi.org/10.13130/2037-3597/13765

Andorno, C. M., & Sordella, S. (2021). Promuovere un'educazione plurilingue parlando a casa di cose di scuola: esiti da una ricerca-azione. [Promoting plurilingual education by talking about school things at home: results from action research]. In S. Caruana; K. Chircop, P. Gauci, & M. Pace (Eds). *Politiche e pratiche per l'educazione linguistica, il multilinguismo e la comunicazione interculturale*, 29-40. Edizioni Ca'Foscari.

Andreou, M., Knopp, E., Bongartz, C., & Tsimpli, I. M. (2015). Character Reference in Greek-German Bilingual Children's Narratives. In L. Roberts, K. McManus, N. Vanek & D. Trenkic (Eds), *EUROSLA Yearbook* 15, 1-40. Amsterdam: John Benjamins.

Andreou, M., Peristeri, E., & Tsimpli, I. M. (2020). Reference maintenance in the narratives of Albanian– Greek and Russian–Greek children with Developmental Language Disorder: A study on crosslinguistic effects. *First Language*, https://doi.org/10.1177/01427237209483

Antoniou, K., Veenstra, A., Kissine, M., & Katsos, N. (2020). How does childhood bilingualism and bi-dialectalism affect the interpretation and processing of pragmatic meanings? *Bilingualism: Language and Cognition*, 23(1), 186–203.

Applebee A. (1978). *The child's concept of a story: Ages 2 to 17*. University of Chicago Press.

Arcodia, G.F., Basciano, B. (2020). Il cinese [Chinese]. In I. Fiorentini, C. Gianollo, & N. Grandi (Eds). *La classe plurilingue*, 159-174, Bologna: Bononia University Press.

Argyri, E., & Sorace, A. (2007). Crosslinguistic influence and language dominance in older bilingual children. *Bilingualism: Language and Cognition*, 10, 79–99.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics*, 24(1), 65–87. DOI: 10.1017/S0022226700011567

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.

Ariel, M. (1991). The function of accessibility in a theory of grammar. *Journal of Pragmatics*, 16(5), 443–464. DOI: 10.1016/0378-2166(91)90136-L

Ariel, M. (1996). Referring expressions and the+/-coreference distinction. In Fretheim, T., Gundel, J.K. (Eds). *Reference and referent accessibility*, 13-35, Amsterdam: John Benjamins.

Ariel, M. (2001). Accessibility theory: An overview. In T. Sanders, J. Schliperoord & W. Spooren (Eds). *Text Representation (Human cognitive processing series)*, 29-87, Amsterdam: John Benjamins.

Arnold, J.E., Zerkle, S.A. (2019). Why do people produce pronouns? Pragmatic selection vs. rational models, *Language, Cognition and Neuroscience*, 1-24. DOI:10.1080/23273798.2019.1636103

Astington, J. W., & Pelletier, J. (2005). Theory of mind, language, and learning in the early years: Developmental origins of school readiness. In B. D. Homer & C. S. Tamis-LeMonda (Eds), *The development of social cognition and communication*, 205–230. Lawrence Erlbaum Associates Publishers.

Atkinson, R.C., Shiffrin R.M. (1968). Human memory: A proposed system and its Control Processes. *Psychology of Learning and Motivation*, 2, 1968, 89-195.

Austin, J (2007). Grammatical interference and the acquisition of ergative case in bilingual children learning Basque and Spanish. *Bilingualism: Language and Cognition* 10 (3), 2007, 315–331. Cambridge University Press DOI:10.1017/S1366728907003094

Avrutin, S., Coopmans, P. (2000). Children who buildb ridges. In S. C. Howell., S.A. Fish, T. Keith-Lucas (Eds). *Proceedings of the Boston University Conference on Language Development 24 (BUCLD24)*, 90-102. Cascadilla Press, Somerville, MA.

Avrutin, S. (1999). *Development of the syntax-discourse interface*. Dordrecht: Kluwer Academic Publishers.

Avrutin, S. (2006). *Weak syntax*. In K. Amundt,, & Y. Grodzinsky (Eds). *Broca's Region*. 49–62. Oxford: Oxford Press.

Avrutin, S. & De Lange, J. (2004*). Children don't read newspapers and yet.... A cross-linguistic comparison of child utterances and newspaper headlines*. Paper presented at the 2nd Lisbon Meeting on Language Acquisition, Lisbon.

Baauw, S., de Roo, E. & Avrutin, S. (2002). Determiner omission in language acquisition and language impairment: Syntactic and discourse factors. In B. Skarabela, S. Fish & A. Do (Eds). *Proceedings of the 26th annual Boston University conference on language development*. Somerville, MA: Cascadilla Press.

Badawi E.S., Hinds M. (1986). *A Dictionary of Egyptian Arabic*. Beirut: Librairie du Liban.

Bagna C., Barni M., Vedovelli M. (2007). Italiano in contatto con lingue immigrate: nuovi modelli e metodi per il neoplurilinguismo in Italia [Italian in contact with immigrant languages: new models and methods for neoplurilingualism in Italy]. In C. Consani, P. Desideri (Eds). *Minoranze Linguistiche. Prospettive, Strumenti, Territori*, 270-290, Roma: Carocci.

Bailini, S. (2016) *La interlingua de lenguas afines. El español de los italianos, el italiano de los españoles* [The interlanguage of related languages. Spanish of the Italian speakers, Italian of the Spanish speakers]. Milan: Hoepli.

Ball, A. F. (2008). Teaching Writing in Culturally Diverse Classrooms. In C.A. MacArthur, S. Graham & J. Fitzgerald (Eds). *Handbook of Writing Research*. 293-310. Guilford Press.

Bamberg, M. (1986). A functional approach for the acquisition of anaphoric relationships, *Linguistics* 24, 227-284.

Bamberg, M.G.W. (1987). *The acquisition of narratives: learning to use language*. Berlin: Mouton de Gruyter.

Bánhidi, Z., Jókay Z. & Szabó, D. (1965), *A textbook of the Hungarian language*, London.

Barton, D.,Hamilton, M. (2000). Literacy Practices. In D. Barton, M. Hamilton, & R. Ivanić (Eds). *Situated Literacies*, 7-15, Routledge.

Bassano, D., Maillochon, I., Korecky-Kröll, K., van Dijk, M., Laaha, S., Dressler, W. U., & van Geert, P. (2011). A comparative and dynamic approach to the development of determiner use in three children acquiring different languages. *First Language*, 31(3), 253–279. DOI: 10.1177/0142723710393102

Bassano, D., Maillochon, I., & Mottet, S. (2008). Noun grammaticalization and determiner use in French children's speech: A gradual development with prosodic and lexical influences. *Journal of Child Language*, 35(2), 403–438. DOI: 10.1017/S0305000907008586

Bassano, D. (2015). The acquisition of nominal determiners. In Serratrice, L., & Allen, S.E.M. (Eds), *The Acquisition of Reference.* 25-49. Amsterdam: John Benjamins Publishing.

Bates, E., & Goodman, J. C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (Ed), *The emergence of language*, 29–79. Hillsdale, NJ: Erlbaum.

Baum, S., & Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Applied Psycholinguistics, 35*(5), 857-894. https://doi.org/10.1017/S0142716414000174

Beacco, J.C., Byram, M., Cavalli, M., Coste, D., Cuenat, M., Goullier, F. & Panthier, J. (2010). *Guide for the development and implementation of curricula for plurilingual and intercultural education*. Strasbourg, Council of Europe. [www.coe.int/T/DG4/Linguistic/Source/Source2010_ForumGeneva/GuideEPI2010_EN.pdf]

Beaver D., Francez I., & Levinson D. (2006). Bad subject: (Non)-canonicality and NP distribution in existentials, M. Gibson, & J. Howell (Eds). *Proceedings of Semantics and Linguistic Theory*, 15, 19-43. Ithaca (NY): CLC Publications.

Becker, L. (2018), *Articles in the world's languages*, PhD thesis, Leipzig University, Leipzig.

Becker,L. (2021). *Articles in the World's Languages*, Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110724424

Belletti, A. (1988). The case of unaccusatives. *Linguistic Inquiry* 19, 1-33.

Belletti, A. & Guasti, M.T. (2015). *The Acquisition of Italian. Morphosyntax and its interfaces in different modes of acquisition*. Amsterdam: John Benjamins.

Bergman, C. (1976). Interference vs. independent development in infant bilingualism. In G. Keller, Teschner, & S. Viera (Eds), *Bilingualism in the bicentennial and beyond*. 86-96, New York: Bilingual Press/Editorial Bilingüe.

Berman, R. (2009). Language development in narrative contexts. In E.L. Bavin (Ed). *Handbook of Child Language*, 1st edn, 354–75. Cambridge University Press.

Berman, R. (2015). Language development and use beyond the sentence. In E.L. Bavin, & L.R. Naigles (Eds). *Handbook of Child Language*, 2nd edn, 458-480. Cambridge University Press.

Berman, R. A., & Slobin, D. I. (1994). *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Hillsdale, NJ: Erlbaum.

Bernardini, P., & Schlytr, S. (2004). Growing syntactic structure and codemixing in the weaker language: The ivy hypothesis. *Bilingualism: Language and Cognition*, 7, 49–69.

Bezuidenhout A. (2019). Joint Reference. In J. Gundel, & B. Abbott (Eds). *The Oxford Handbook of Reference*, Oxford Handbooks.

Biazzi, M. (2018). Il comportamento linguistico [Linguistic Behavior]. In M. Chini, & C. Andorno (Eds). *Repertori e usi linguistici nell'immigrazione*, 119-160, Milano: Franco Angeli.

Bock, K.J. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387.

Bohman, T., Bedore, L., Pe.a, Mendez-Perez, A., & Gillam, R. (2010). What you hear and what you say: Language performance in Spanish-English bilinguals. *International Journal of Bilingual Education and Bilingualism*, 13, 325–344.

Bohnacker, U. & Gagarina, N. (2020). *Developing Narrative Comprehension*. Amsterdam: John Benjamin. DOI: https://doi.org/10.1075/sibil.61

Bonifacio, S., & Hvastja Stefani, L. (2010). *L'intervento precoce nel ritardo di linguaggio: il modello INTERACT per il bambino parlatore tardivo* [Early intervention in language delay: the INTERACT model for the late speaking child]. Milano: Franco Angeli.

Borg, C., Cardona, M., & Caruana, S. (2013). *Social class, language and power. 'Letter to a teacher': Lorenzo Milani and the School of Barbiana*. Rotterdam Boston & Taipei: Sense editors

Bosch, J. E., & Unsworth, S. (2020). Cross-linguistic influence in word order: effects of age, dominance and surface overlap. *Linguistic Approaches to Bilingualism*, 1–34. https://doi.org/10.1075/lab.18103.bos

Bredart, S. (1987). Is the quantity maxim strategy child-specific? *Journal of Child Language*, 14(1), 175-7. DOI: 10.1017/s0305000900012800

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press. DOI: 10.4159/harvard.9780674732469

Brugè, L. (1996). Demonstrative movement in Spanish: A comparative approach. *Working papers in Linguistics*, 6(1), 1-53.

Brugger,G. (1993). Generic interpretations and expletive determiner. *University of Venice Working Papers in Linguistics,* 3(1),1.

Brustad K. (2000), *The syntax of spoken Arabic: a comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Washington, D.C: Georgetown University Press

Bruner, J. (1986). *Actual minds, possible worlds.* Harvard University Press.

Bruner, J. (1990). *Acts of meaning*. Cambdridge (MA)-London: Harvard University Press.

Bruner, J. S. (2002). *Making stories: Law, literature, life*. New York: Farrar, Strauss and Giroux. Tr. It. La fabbrica delle storie. Diritto, letteratura, vita. Laterza, Roma – Bari: 2002.

Bryant, J.B. (2015). Pragmatic development. In E-L. Bavin, & L.R. Naigles, (Eds). *Handbook of Child Language*, 2nd edn, 438-457. Cambridge University Press.

Butt J. & Benjamin C. (1994). *A new reference grammar of Modern Spanish*, London: Arnold.
Calabrese, A. (1988). I dimostrativi: pronomi e aggettivi [Demonstratives: pronouns and adjectives]. In L. Renzi, G. Salvi, & A. Cardinaletti (Eds). *Grande grammatica italiana di consultazione*3., 1st vol., 617-631. Bologna: il Mulino.

Calvi, M.V., Mapelli, G., & Bonomi, M. (2011). *Lingua, identità e immigrazione. Prospettive interdisciplinari* [Language, identity and immigration. Interdisciplinary perspectives]. Milano:

Franco Angeli.

Cairney, T.H. (2009). Home Literacy Practices and Mainstream Schooling: A Theoretical Understanding of the Field. In Li G. (Ed). *Multicultural Families, Home Literacies, and Mainstream Schooling*, 3-24, Information Age Publishing.

Camaioni, L. (2022). *La Teoria della mente. Origini, sviluppo e patologia* [Theory of Mind. Origins, development and impairment]. Bari: Editori Laterza.

Candelier M., Camilleri-Grima A., Castellotti V., de Pietro J.-F., Lörincz I., Meißner F.- J., Schröder-Sura A., Noguerol A., Molinié M. (2011). *Le CARAP - Un Cadre de Référence pour les Approches Plurielles des langues et des cultures. Compétences et ressources.* [The FREPA-A Framework of Reference for Pluralistic Approaches. Competences and resources]. Conseil de l'Europe. Strasbourg - Graz: http://carap.ecml.at/Handbook/tabid/2602/language/fr-FR/Default.aspx

Cantone, K., Kupisch, T., Müller, N., & Schmitz, K. (2008). Rethinking language dominance in bilingual children. *Linguistische Berichte*, 215, 307–343.

Carbonara, V., & Scibetta, A. (2020). Integrating translanguaging pedagogy into Italian primary schools: implications for language practices and children's empowerment, *International Journal of Bilingual Education and Bilingualism*, DOI: 10.1080/13670050.2020.1742648

Carbonara, V., Scibetta, A., & Torregrossa, J. (2023). The benefits of multilingual pedagogies for multilingual children's narrative abilities, *International Journal of Multilingualism*. DOI: 10.1080/14790718.2022.2148675

Cardinaletti, A., & Giusti, G. (2016). The syntax of the Italian indefinite determiner dei. *LINGUA*, 181, 58-80.

Cardinaletti, A., & Giusti, G. (2018). Indefinite determiners. Variation and Optionality in Italoromance. In R. D'Alessandro, & D. Pescarini (Eds). *Advances in Italian Dialectology. Sketches of Italo-Romance Grammars*, 1, 135-161. Amsterdam: Brill.

Carlson, G. (2012). 47. Genericity. In K. Heusinger, C. Maienborn,, & P. Portner (Eds). *Volume 2*, 1153-1185. Berlin, Boston: De Gruyter
Mouton. https://doi.org/10.1515/9783110255072.1153

Carmiol, A., M., & Sparks, A. (2014). Narrative Development across Cultural Contexts: Finding the Pragmatic in Parent-Child Reminiscing. In D. Matthews (Ed). *Pragmatic Development in First Language Acquisition*. 279-294, Amsterdam: John Benjamin.

Caruana, S., & Scaglione, S. (2014). 'Superdiverse' school populations in southern Europe: reflections on language use and suggestions for learning strategies. *International Journal of Cross-cultural Studies and Environmental Communication* (Special issue: "Dialogue-driven Change in the Public Sphere"), 1(1), 63-73.

Caselli, M. C., Leonard, L. B., Volterra, V. & Campagnoli, M. G. (1993). Toward mastery of Italian morphology: A cross- sectional study. *Journal of Child Language*, 20, 377–393.

Caso, R. (2017). Bambini e malattia. Le storie che curano [Children and Illness. Healing stories]. In A. Bondioli, & D. Savio (Eds). *Crescere bambini. Immagini di infanzia in educazione e formazione degli adulti*, 113-126, Bergamo: Junior Spaggiari.

Castilla-Earls, A. P., Pérez-Leroux, A. T., & Auza, A. (2021). Elicited vs. spontaneous language as methods for the assessment of grammatical development: The DEME assessment tool. *Revista de Logopedia, Audiología y Foniatría,* 41(4), 164-171.https://doi.org/10.1016/j.rlfa.2021.02.002.

Chafe, W.L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, C.N. (Ed). *Subject and Topic*, 25–55. New York: Academic Press.

Chandler, M.C., Gerde, H.K, Bowles, R.P, McRoy, K.Z., Pontifex, M.B., Bingham, G.E., (2021). Self-regulation moderates the relationship between fine motor skills and writing in early childhood, *Early Childhood Research Quarterly*, 57, 239-250.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological review*, 113(2), 234.

Chao Y.R. (1968), *A Grammar of Spoken Chinese*, Berkeley and Los Angeles: University of California Press.

Chen, L., & Lei, J. (2013). The production of referring expressions in oral narratives of Chinese–English bilingual speakers and monolingual peers. Child Language Teaching and Therapy, 29(1), 41-55. https://doi.org/10.1177/0265659012459527

Chen L., & Pan, N. (2009). Development of English referring expressions in the narratives of Chinese–English bilinguals. *Bilingualism: Language and Cognition* 12: 429–445.

Chen P. (2003). Indefinite determiner introducing definite referent: A special use of 'yi 'one' + 'classifier' in Chinese*, Lingua,* 113, 1169-1184.

Chen P. (2004). Identifiability and definiteness in Chinese, *Linguistics*, 42-6: 1129-1184.

Chen P. (2009). Aspects of referentiality, *Journal of Pragmatics*, 41: 1657-1674.

Chen P. (2015). Referentiality and Definiteness in Chinese. In W.S.Y. Wang, & C. Sun (Eds). *The Oxford handbook of Chinese linguistics*, 404-413. Oxford: Oxford University Press.

Cheng L.L., Sybesma R. (1999). Bare and not-so-bare nouns and the structure of NP, *Linguistic Inquiry*, 30, 509-542.

Chiapedi, N. (2010). L'articolo italiano nell'interlingua di apprendenti sinofoni: problematiche acquisizionali e considerazioni glottodidattiche [The Italian article in the interlanguage of Chinese-speaking learners: acquisitional problems and teaching considerations], *Italiano LinguaDue*, 2.2, 53-74.

Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics*, 6(4), 339–405. DOI: 10.1023/A:1008324218506

Chierchia, G., Guasti, M. T. & Gualmini, A. (1999). Nouns and articles in child grammar and the syntax/semantics map. Presentation given at GALA, Potsdam, Germany.

Chini, M. (1995). *Genere grammaticale e acqusizione. Aspetti della morfologia nominale in italiano L2*. Milano: FrancoAngeli

Chini, M., & Andorno, C. (Eds). (2018). *Repertori e usi linguistici nell'immigrazione* [Linguistic repertoire and uses in immigration]. Milano: Franco Angeli.

Chondrogianni, V., Marinis, T. (2015). Production of definite and indefinite articles in typically developing English-speaking children and children with SLI, *Lingua*, 155, 9–2.

Chondrogianni, V., Marinis, T., Edwards, S., & Blom, E. (2015). Production and on-line comprehension of definite articles and clitic pronouns by Greek sequential bilingual children and monolingual children with specific language impairment, *Applied Psycholinguistics*, 36(5), 1155–1191.

Clark, H.H., & Marshall, C.R. (1981). Definite reference and mutual knowledge. In A.K. Joshi, B.L. Webber, & I.A. Sag (Eds). *Elements of discourse understanding*, 10–63. Cambridge: Cambridge University Press.

Comrie, B. (1978). Definite direct objects and referent identification, *Pragmatics Microfiche*, 3.1.D3.

Consensus Conference (2007). *Disturbi Evolutivi Specifici di Apprendimento* [Specific Developmental Learning Disorders]. Milan, 26th January 2007. https://www.miur.gov.it/documents/20182/0/Raccomandazioni+per+la+pratica+clinica+definite+con+il+metodo+della+Consensus+Conference.pdf/b504867d-ecba-4f72-9589-2e46a603bd53?version=1.0&t=1495725551691

Contento, S., Bellocchi, S., Bonifacci, P. (2013). *BaBIL- Prove per la valutazione delle competenze verbali e non-verbali in Bambini Bilingui* [BaBIL- Tests for the assessment of verbal and non-verbal skills in bilingual children], Firenze: Giunti O.S.

Costa, A. (2004). Speech production in bilinguals. In T.K.Bhatia & W. C. Ritchie (Eds). *The handbook of bilingualism*, 201–223. Malden, MA: Blackwell Publishing.

Cotter W., de Jong R. (2019). Regional variation. In E. Al-Wer, & U. Horesh (Eds). *The Routledge Handbook of Arabic Sociolinguistics*, 47-62. Abingdon, Oxon, New York: Routledge.

Crisma, P., & Tomasutti, E. (2000). Phonological effects on article omission in the acquisition of Italian. In S.C. Howell, S. Fish, & T. Keith-Lucas (Eds). *Proceedings of the 24th Boston University Conference on Language Development*, T. 220-231. Somerville: Cascadilla Press.

Croft W. (1990). *Typology and universals*. Cambridge, Cambridge University Press.

Cummins, G. M. (1998). Definiteness in Czech. *Studies in Language*, 22(3), 567–596.

Cummins, J. (1979). Linguistic Interdependence and the Educational Development of Bilingual Children. *Review of Educational Research*, 49, 222-5.
Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Multilingual Matters.

Cummins, J. (2005). A Proposal for Action: Strategies for Recognizing Heritage Language Competence as a Learning Resource within the Mainstream Classroom. *The Modern Language Journal*, *89*(4), 585–592. http://www.jstor.org/stable/3588628

Cummins, J. (2017). Teaching Minoritized Students: Are Additive Approaches Legitimate? *Harvard Educational Review,* 87(3), 404-426.
https://www.cpatrickproctor.com/uploads/9/0/9/1/9091150/cummins_2017.pdf

Cummins, J., & Early, M. (2011). *Identity texts: The collaborative creation of power in multilingual schools*. Trentham Books Ltd.

Daloiso, M. (2015). La rilevazione sistematica della competenza comunicativa dei bambini esposti alla lingua straniera. Una ricercar nella scuola dell'infanzia [The systematic survey of the communicative competence in children exposed to a foreign language. Research in kindergartens], *EL. LE*, 4(3), 381-410. DOI: 10.14277/2280-6792/ELLE-4-3-15-1

Dayal, V. (2004). Number Marking and (In)definiteness in Kind Terms. *Linguistics and Philosophy,* 27(4), 393-450.

Dayal, V. (2018). (In)definiteness without articles: Diagnosis, Analysis, Implications. In Sharma G. & Bhatt R. (Eds). *Trends in Hindi Linguistics*, Berlin: Mouton de Gruyter.

De Bot, K. (1992). A Bilingual Production Model: Levelt's 'Speaking' Model Adapted. *Applied Linguistics*, 13, 1-24.

De Cat, C. (2011). Information tracking and encoding in early L1: Linguistic competence vs. cognitive limitations. *Journal of Child Language*, 38(4), 828–860. DOI: 10.1017/S030500091000036X

De Cat, C. (2013). Egocentric definiteness errors and perspective evaluation in preschool children. *Journal of Pragmatics*, 56, 58–69. DOI: 10.1016/j.pragma.2012.08.002

De Cat, C. (2015). The cognitive underpinnings of referential abilities. In Serratrice, L., & Allen, S.E.M. (Eds). *The Acquisition of Reference.* 25-49. Amsterdam: John Benjamins Publishing.

De Cat C., Gusnanto A. & Serratrice L. (2017). Identifying a threshold for the executive function advantage in bilingual children. *Studies in Second Language Acquisition*, 40, 119-151.

De Cat, C. (2021). Socio-economic status as a proxy for input quality in bilingual children? *Applied Psycholinguistics*, 42, 301–324.

De Cat, C. (2022). Opportunities and Challenges in the Analysis of the Multilingual Assessment Instrument for Narratives (MAIN). *First Language*,42(2), 321-329.
DOI: 10.1177/01427237211064695

De Cat, C., & Unsworth, S. (2023). So many variables, but what causes what? *Journal of Child Language*, 50(4), 832-836. DOI: 10.1017/S0305000923000107.
De Lange, J., Vasic, N., Avrutin, S. (2009). Reading between the (head)lines: A processing account of article omissions in newspaper headlines and child speech. *Lingua,* 119 (2009), 1523–1540. DOI: 10.1016/j.lingua.2008.04.005
De Houwer, A. (1990). *The acquisition of two languages from birth: A case study.* Cambridge University Press. https://doi.org/10.1017/CBO9780511519789

De Houwer, A. (2020). Harmonious Bilingualism: Well- being for Families in Bilingual Settings. In A.C. Schalley, S.A Eisenchlas (Eds). *Handbook of Home Language Maintenance and Development*. Berlin/Boston: De Gruyter.

De Houwer, A. (2021). *Bilingual development in childhood.* Cambridge University Press. https://doi.org/10.1017/9781108866002

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, 23, 517– 542.

Dell, G. S., Reed, K. D., Adams, D. R. & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1355–1367.

Della Puppa, F. (2007). Dall'analisi del contatto arabo-italiano L2 alle implicazioni glottodidattiche [From the analysis of Arabic-Italian L2 contact to language teaching implications], *Studi di glottodidattica*, 3, 22-40.

Demuth, K. (2011). The acquisition of phonology. In J. Goldsmith, J. Riggle, & A. Yu (Eds). *The handbook of phonological theory*, 29-79. Malden, MA: Blackwell.

Despic M. (2019). On kinds and anaphoricity in languages without definite articles. In A. Aguilar-Guevara, J. Pozas Loyo, V. Vázquez-Rojas Maldonado (Eds). *Definiteness across languages*, 259:292. Berlin: Language Science Press.

Dijkstra, T., & van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision*, Bilingualism: Language and Cognition, 5(*3), 175-197. https://doi.org/10.1017/S1366728902003012

d.m. 16/11/2012, n. 256. *Regolamento recante indicazioni nazionali per il curricolo della scuola dell'infanzia e del primo ciclo d'istruzione* [Regulation containing national indications for the curriculum of kindergartens and the first cycle of education].
Donnellan, K.S. (1966). Reference and definite descriptions, *Philosophical Review*, 75, 281-394.

Döpke, S. (1998). Competing language structures: The acquisition of verb placement by bilingual German-English children, *Journal of Child Language*, 25, 555–584.

Dryer, M. S. (2013). Definite Articles. In M.S. Dryer, & M. Haspelmath (Eds). *The World Atlas of Language Structures*. Leipzig: Max Planck Institute for Evolutionary Anthropology http://wals.info/chapter/37. Accessed on 2023-08-24.

Du Bois, J. W. (1980). Beyond definiteness: the trace of identity in discourse. In W. Chafe (Ed). *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*, Norwood, NJ: Albex Publishing Corporation: 203–274.

Emslie, H., Stevenson, R. (1981). Pre-school children's use of the articles in definite and indefinite referring expressions, *Journal of Child Language* 8, 313-328.

Enç,M. (1991). The semantics of specificity. *Linguistic Inquiry* 22, 1–25.

Epstein, R. (2002). The definite article, accessibility, and the construction of discourse referents, *Cognitive Linguistics*,12(4), 333-378.

Espinal, M. T. (2010). Bare nominals in Catalan and Spanish. Their structure and meaning, *Lingua*, 120 (4), 984-1009.

Espinal M.T., Cyrino S. (2017). The definite article in Romance expletives and long weak definites, *Glossa: a journal of general linguistics,* 2(1), 23, 1-26.

Evans, N. & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science, *Behavioral and Brain Sciences*, 32(5):429-448.

Evers, S. (2020). *To 'the' or not to 'the', Cross-Linguistic Correlations Between Existing Morphosyntax and the Emergence of Definite Articles*. PhD thesis, Buffalo, University, Buffalo.

Ezeizabarrena, M.J., & Garcia del Real, I. (2020). The Spanish adaptation of MAIN, *ZAS Papers in Linguistics* 64, 211-220.

Fassi Fehri, A. (2005). Verbal and nominal parallelism and extensions, Reports *& Documents* no 8. 1-22, Rabat. IERA.

Favaro, G., Demaerschalk, B., Enea, S., Nuzzo, E., Pallotti, G. (2012). *Dare parole al mondo. L'italiano dei bambini stranieri* [Giving words to the world. The Italian of foreign children]. Bergamo: Edizioni Junior.

Fernández Fuertes, R., & Liceras, J. (2010). Copula omission in the English developing grammar of English/Spanish bilingual children, *International Journal of Bilingual Education and Bilingualism*, 13, 525–551.

Ferreira, V., & Dell, G. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production, *Cognitive Psychology*, 40, 296–340.

Ferretti, L.K., & Bub, K.L. (2017). Family Routines and School Readiness During the Transition to Kindergarten, *Early Education and Development*, 28(1), 59-77. DOI: 10.1080/10409289.2016.1195671

Fichman, S., & Altman, C. (2019). Referential cohesion in the narratives of bilingual and monolingual children with typically developing language and with specific language impairment, *Journal of Speech, Language, and Hearing Research*, *62*(1), 123-142. https://doi.org/10.1044/2018_JSLHR-L-18-0054

Fichman, S., Walters, J., Melamed, R., & Altman, C. (2020). Reference to characters in narratives of Russian-Hebrew bilingual and Russian and Hebrew monolingual children with Developmental Language Disorder and typical language development, *First Language*, *1*, 29.

Fillmore, C. (1982). Frame Semantics. In Linguistic society of Korea (Ed). *Linguistic in the morning calm*. Selected paper from SICOL-1981. Seoul: Hanshin Publishing Company.

Finnstedt, L. (2013). *Referentiell kohesion och temporal förankring i barns narrativer: Jämförelse mellan enspråkiga svensktalande och flerspråkiga engelsk-svensktalande barn* [Referential cohesion and temporal anchoring in children's narratives: A comparison of monolingual Swedish-speaking and multilingual English-Swedish-speaking children] [Unpublished MA thesis]. Uppsala University, Uppsala.

Fiorentini I., Gianollo, C., (2018). *Evviva le lingue del mondo! Questionario sociolinguistico* [World Languages! Sociolinguistic Questionnaire], Progetto "La classe plurilingue. Ricerca sulla complessità linguistica per una didattica inclusiva". Bologna, Università di Bologna e Fondazione Alsos.

Fiorentini, I., Gianollo, C., Grandi, N. (Eds). (2020). *La classe plurilingue* [The plurilingual classroom]. Bologna: Bononia University Press.

Flynn, E. E. (2019). Story circles: Supporting boys' social and emotional worlds in school. In K.J. Kerry-Moran, & J. Aerila (Eds). *Story in Children's Lives: Contributions of the Narrative Mode to Early Childhood Development, Literacy, and Learning, Educating the Young Child* 16, 97-112. https://doi.org/10.1007/978-3-030-19266-2_6

Fodor, J. & Sag, I. (1982). Referential and quantificational indefinites. *Linguistics & Philosophy* 5, 355–398.

Foroodi-Nejad, F., & Paradis, J. (2009). Crosslinguistic transfer in the acquisition of compound words in Persian-English bilinguals, *Bilingualism: Language and Cognition*, 12, 411–427. https://doi.org/10.1017/S1366728909990241

Frege, G. (1892). *Über Sinn und Bedeutung. Zeitschrift für Philosophie und Philosophische Kritik,* 100: 22–50.

Friedman, N., Miyake, A., Corley, R., Young, S., DeFries, J., & Hewitt, J. (2006). Not all executive functions are related to intelligence. *Psychological Science*, 17, 172–179.DOI: 10.1111/j.1467-9280.2006. 01681.x.

Fukumura, K., & Van Gompel, R. (2012). Producing pronouns and definite noun phrases: Do speakers use the addressee's discourse model? *Cognitive Science*, 36, 1289–1311. DOI: 10.1111/j.1551-6709.2012. 01255.x

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2012). Multilingual Assessment Instrument for Narratives (MAIN). *ZAS Papers in Linguistics*, 56. Berlin: Zentrum für Allgemeine Sprachwissenschaft.

Gagarina, N., Klop, D., Tsimpli, I., & Walters, J. (2016). Narrative abilities in bilingual children. *Applied Psycholinguistics*, *37*(1), 11-17. https://doi.org/10.1017/S0142716415000399

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Bohnacker, U., & Walters, J. (2019a), MAIN: Multilingual Assessment Instrument for Narratives – Revised, *ZAS Papers in Linguistics*, 63, 1–36.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Bohnacker, U. & Walters, J. (2019b). MAIN: Multilingual Assessment Instrument for Narratives – Revised. Materials for use. Arabic (Lebanese version). Translated and adapted by Haddad, R. *ZAS Papers in Linguistics, 63.*

Gagarina, N., Musan, R. (2020). Introduction. In Gagarina, N., Musan, R. (Eds*), Referential and Relational Discourse Coherence in Adults and Children*. 139-170. Berlin: Mouton De Gruyter.

Gagarina, N., & Bohnacker, U. (2022). A new perspective on referentiality in elicited narratives: Introduction to the Special Issue. *First Language*, *42*(2), 171-190.

Gagliardi, G. (2019). *Le competenze narrative: sviluppo fisiologico e patologico* [Narrative skills: physiological and pathological development]. Presentation given at Giornata di studio "Le abilità narrative nel bambino", Azienda USL, Toscana Centro.

Galante, A., Chiras, M., dela Cruz, J. W. N., & Zeaiter, L. F. (2022). *Plurilingual guide: Implementing critical plurilingual pedagogy in language education*. Plurilingual Lab Publishing. https://www.mcgill.ca/plurilinguallab/files/plurilinguallab/plurilingual_guide.pdf

Gámez, P. B., & Vasilyeva, M. (2020). Shared syntactic representations in balanced bilinguals: Crosslinguistic priming with and without verb overlap. *Language Learning and Development, 16*(1), 89-106. https://doi.org/10.1080/15475441.2019.1689822

García, O., Flores, N., & Spotti, M. (Eds). (2017). *The Oxford handbook of language and society*. Oxford University Press.

García, O., & Kleyn, T. (Eds). (2016). *Translanguaging with multilingual students: Learning from classroom moments*. Routledge.

Garcìa, O., Wei, L. (2014). Language, Bilingualism and Education. In O. Garcìa, L. Wei (Eds). *Translanguaging: Language, Bilingualism and Education*, 46-62. London: Palgrave Macmillan.

Garon, N., Bryson, S., & Smith, I. (2008). Executive function in preschoolers: A review using an integrative framework, *Psychological Bulletin*, 134, 31–60. DOI: 10.1037/0033-2909.134.1.31

Gavarró, A. (2003). Economy and word order patterns in bilingual English-Dutch acquisition, *Bilingualism: language and cognition*, 6(1), 69-79.

Gavarró, A., Pérez-Leroux, A. T., & Roeper, T. (2006). Definite and bare noun contrasts in child Catalan. In V. Torrens, & L. Escobar (Eds). *The acquisition of syntax in Romance languages*, 51-68. Amsterdam: John Benjamin.

Gawlitzek-Maiwald, I., & Tracy, R. (1996). Bilingual bootstrapping, *Linguistics*, 34, 901–926. https://doi.org/10.1515/ling.1996.34.5.901

Genesee, F. (1989). Early bilingual development: One language or two?, *Journal of Child Language,* 16, 161–179. https://doi.org/10.1017/S0305000900013490

Genesee, F., Boivin, I., & Nicoladis, E. (1996). Talking with strangers: A study of bilingual children's communicative competence. *Applied Psycholinguistics*, 17(4), 427–442. doi:10.1017/S0142716400008183
Gerards, D. P., & Stark, E. (2022). Non-Maximal Definites in Romance. *Isogloss. Open Journal of Romance Linguistics*, *8*(5), 1-32.

Gerken, L. A. (1994). A metrical template account of children's weak syllable omissions from multisyllabic words, *Journal of Child Language*, 21(3), 565–584. DOI: 10.1017/S0305000900009466

Gerken, L. A. (1996). Prosodic structure in young children's language production, *Language*, 72(4), 683–712. DOI: 10.2307/416099

Giacalone Ramat, A. (Ed). (2017). *Verso l'italiano. Percorsi e strategie di acquisizione*. [Towards Italian. Acquisition paths and strategies]. Roma: Carocci.

Gianollo, C., Silvestri, G. (2022). *Key Topics in Semantics: Presupposition, Anaphora, (In)definite Nominal Phrases, Deixis, Tense and Aspect, Negation*. In a. Ledgeway, A. Maiden (Eds). *The Cambridge Handbook of Romance Linguistics*, 695-727. Cambridge: Cambridge University Press.

GISCEL (Gruppo di Intervento e Studio nel Campo dell'Educazione Linguistica). (1975). *Dieci Tesi per un'educazione linguistica democratica* [Ten theses for a democratic linguistic education]. https://giscel.it/dieci-tesi-per-leducazione-linguistica-democratica/.

Giusti, G. (1994). Enclitic articles and double definiteness: A comparative analysis of nominal structure in Romance and Germanic, *The Linguistic Review*, 11, 241-256.

Giusti, G., Gozzi, R. (2006). The acquisition of determiners: Evidence for the Full Competence Hypothesis, Language acquisition and development. In A. Belletti (Ed). *Proceedings of GALA 2005*, 232-237. Cambridge: Cambridge Scholars Publishing.

Givón, T. (1981). On the development of the numeral 'one' as an indefinite marker, *Folia Linguistica Historica.*2, 35–53.

Givón, T. (1983). Topic Continuity in Discourse. An Introduction. In Givón, T. (ed.). *Topic Continuity in Discourse. A Quantitative Cross-Language Study*. 1–41. Amsterdam, Philadelphia: Benjamins.

Givón, T. (1984). *Syntax: A Functional-Typological Introduction*. 1st vol. Amsterdam: J. Benjamins.

Givón, T. (1990). *Syntax: A Functional-Typological Introduction*. 2nd vol. Amsterdam: J. Benjamins.

Givón, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics*, 30(5), 5-55.

Givón, T. (2017). The story of zero. *The Story of Zero*, 1-430. Amsterdam: John Benjamin.

González Luna, A.M., & Sagi-Vela, A. (2020). Varietà di spagnolo dell'Ispanoamerica. In I. Fiorentini, C.Gianollo, & N. Grandi (Eds). *La classe plurilingue*, 159-174, Bologna: Bononia University Press.

González, N.; Moll, L.C. & Amanti, K. *Funds of Knowledge: Theorizing Practices in Households, Communities, and Classrooms*. Mahwah NJ, Lawrence Erlbaum Associates.

Gottardo, A., B. Yan, L. Siegel, and L. Wade-Wooley. 2001. Factors related to English reading performance in children with Chinese as a first language: More evidence of cross-language transfer of phonological processing, *Journal of Educational Psychology* 93: 530 42.

Gottardo, A., Al-Janaideh, R., Paradis, J., Soto-Corominas, A., Chen, X., & Amin, N. (2023). Age, Experience and Language and Literacy Skills in English-Arabic Speaking Syrian Refugees, *Languages*, 8(2), 101.

Greenberg, J.H. (1978). How does a language acquire gender markers? In Greenberg, J.H., Ferguson, C.A., & Moravcsik, E.A. (Eds). *Universals of human language*, 3rd vol., 47–82. Stanford: Stanford University Press.

Grice, H.P. (1975). Logic and conversation. In Cole, P., & Morgan J.L. (eds). *Syntax and semantics*, 3rd vol., 41–58. New York: Academic Press.

Grigoroglou, M., & Papafragou, A. (2019). Interactive contexts increase informativeness in children's referential communication. *Developmental psychology*, *55*(5), 951.

Groba, A., De Houwer, A., Mehnert, J., Rossi, S., & Obrig, H. (2018). Bilingual and monolingual children process pragmatic cues differently when learning novel adjectives, *Bilingualism: Language and Cognition*, *21*(2), 384-402.2018

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Harvard University Press.

Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modeling the local discourse. *Computational Linguistics*, 21, 203–225.

Guardiano, C. (2012). I dimostrativi e il dominio nominale. Note per una lettura parametrica [Demonstratives and the nominal domain. Notes for parametric reading]. In V. Bambini, I. Ricci, P.M. Bertinetto (Eds). *Linguaggio e cervello-Semantica/Language and the brain-Semantics*. 2nd vol. Rome: Bulzoni.

Guasti, M. T., De Lange, J., Gavarró, A. & Caprin, C. (2004). Article omission: Across child languages and across special registers. In Van Kampen, J. & Baauw, S. (Eds). *Proceedings of Gala 2003*, 1st vol., 199–210. Utrecht: Lot Occasional Series.

Guasti, M. T., De Lange J., Gavarró, A., & Caprin, C. (2008). Article omission across child languages, *Language Acquisition*, 15(2), 89–119. DOI: 10.1080/10489220801937346

Guerriero, S., Oshima-Takane, Y., & Kuriyama, Y. (2006). The development of referential choice in English and Japanese: A discourse-pragmatic perspective*, Journal of Child Language, 33,823–857.*

Gülzow, I., & Gagarina, N. (2007). *Frequency effects in language acquisition: Defining the limits of frequency as an explanatory concept*. Berlin: Mouton de Gruyter.

Gundel, J. (2009). Children's use of referring expressions: What can it tell us about theory of mind? *Cognitive Critique*, 1, 73–100

Gundel, J. (2011). Child language, theory of mind, and the role of procedural markers in identifying referents of nominal expressions. In V. Escandell-Vidal, M. Leonetti, & A. Ahern (Eds), *Procedural meaning: Problems and perspectives*, 205-31. Bingley: Emerald Group Publishing Limited

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. Language, 69(02), 274–307. DOI: 10.2307/416535

Gundel, J. K., Hedberg, N., & Zacharski, R. (2001). Definite descriptions and cognitive status in English: Why accommodation is unnecessary. *English Language & Linguistics*, *5*(2), 273-295.

Gundel, J. K., Ntelitheos, D., & Kowalsky, M. (2007). Children's use of referring expressions: some implications for theory of mind. *ZAS Papers in Linguistics*, *48*, 1-21.
Gundel, J. K., & Johnson, K. (2013). Children's use of referring expressions in spontaneous discourse: Implications for theory of mind development, *Journal of Pragmatics*, 56, 43–57. DOI: 10.1016/j.pragma.2013.04.003

Gundel J., Abbott B. (2019). Introduction. In Gundel J., Abbott B. (Eds), *The Oxford Handbook of Reference*. Oxford Handbooks, Oxford University Press.

Hacohen, A., & Schaeffer, J. (2007). Subject realization in early Hebrew/English bilingual acquisition: The role of crosslinguistic influence, *Bilingualism: Language and Cognition*, 10, 333–344.

Hadley, P., Rispoli, & Hsu, N. (2016). Toddlers' verb lexicon diversity and grammatical outcomes. Language, Speech, and Hearing Services in Schools, 4 (1), 44-58. DOI: 10.1044/2015_LSHSS-15-0018

Harning, K.E. (1980). The Analytical Genitive in the Modern Arabic Dialects. *Orientalia Gothoburgensia*, 5. Stockholm.

Harrell, R. S. (2004). *A short reference grammar of Moroccan Arabic: With audio CD*. Georgetown University Press.

Hartsuiker, R. J., Beerts, S., Loncke, M., Desmet, T., & Bernolet, S. (2016). Cross-linguistic structural priming in multilinguals: Further evidence for shared syntax. *Journal of memory and language*, 90, 14-30.

Hartsuiker, R. J., & Bernolet, S. (2017). The development of shared syntax in second language learning. *Bilingualism: Language and Cognition*, 20(2), 219-234. https://doi.org/10.1017/S1366728915000164

Hartsuiker, R., Pickering, M., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish/English bilinguals. *Psychological Science*, 15, 409–414. https://doi.org/10.1111/j.0956-7976.2004.00693.x

Haspelmath, M. (1997). *Indefinite pronouns*. Oxford: Oxford University Press.

Hauser-Grüdl, N., Arencibia Guerra, L., Witzmann, F., Leray, E., & Muller, N. (2010). Crosslinguistic influence in bilingual children: Can input frequency account for it? *Lingua*, 120, 2638–2650.

Hawkins, J.A. (1978). *Definiteness and indefiniteness: a study in reference and grammaticality prediction*. London: Croom Helm.

Haznedar, B. (2007). Crosslinguistic influence in Turkish-English bilingual first language acquisition: The overuse of subjects in Turkish. In A. Belikova, L. Meroni, & M. Umeda (Eds). *Proceedings of the 2nd Conference on Generative Approaches to Second Language Acquisition North America (GALANA)*, 124-134. Somerville, MA: Cascadilla Press.

Heim, I.R. (1988). *The semantics of definite and indefinite noun phrases*. New York: Garland.

Heine, B. (1997). *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Doctoral dissertation. University of Massachusetts.

Henderson, A. T., & Mapp, K. L. (2002). *A New Wave of Evidence: The Impact of School, Family, and Community Connections on Student Achievement*. Annual Synthesis, 2002.

Hendriks, P. (2016). Cognitive modeling of individual variation in reference production and comprehension, *Frontiers in Psychology*, 7, 1–17. 10.3389/fpsyg.2016.00506 https://doi.org/10.3389/fpsyg.2016.00506

Hervé, C., Serratrice, L., & Corley, M. (2016). Dislocations in French-English bilingual children: An elicitation study. *Bilingualism: Language and Cognition*, 19(5), 987-1000. https://doi.org/10.1017/S1366728915000401

Hervé, C., & Serratrice, L. (2018). The development of determiners in the context of French–English bilingualism: a study of cross-linguistic influence. *Journal of child language*, *45*(3), 767-787.

Heycock, C., & Zamparelli, R. (2000). Friends and colleagues: Plurality and NP-coordination. In *North East Linguistics Society*(Vol. 30, No. 1, p. 25).

Heycock, C., & Zamparelli, R. (2003). Coordinated bare definites. *Linguistic Inquiry*, *34*(3), 443-469.
Heycock, C., & Zamparelli, R. (2005). Friends and colleagues: Plurality, coordination, and the structure of DP. *Natural language semantics*, *13*, 201-270.

Hickmann, M. (2003a). Coherence, cohesion, and context in narrative development: Some comparative perspectives. In S. Strömqvist & L. Verhoeven (Eds), Relating events in narrative: Typological and contextual perspectives, 281-306. Hillsdale, NJ: Erlbaum.

Hickmann, M. (2003b). *Children's discourse: Person, space and time across languages*. Cambridge: Cambridge University Press.

Hickmann, M., Kail, M. and Roland, F. (1995). Cohesive anaphoric relations in French children's narratives as a function of mutual knowledge. *First Language* 15, 277–300.
Hickmann, M., Hendriks, H. & Roland, F. (1998). Référence spatiale dans les récits d'enfants. *Langue Française*, Numéro spécial: Acquisition du français langue maternelle 118, 104-126.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child development*, *74*(5), 1368-1378.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental review*, *26*(1), 55-88.

Hsin, L., Legendre, G., & Omaki, A. (2013). Priming crosslinguistic interference in Spanish-English bilingual children. In S. Baiz, N. Goldman, & R. Hawkes (Eds). *Proceedings of the 37th Annual Boston University Conference on Language Development*, 165-177. Cascadilla Press.

Hulk, A. (2004). The acquisition of the French DP in a bilingual context. In P. Prévost & J. Paradis (Eds), *The acquisition of French in different contexts: Focus on functional categories* 243-274. Amsterdam: John Benjamins. DOI: 10.1075/lald.32.12hul

Hulk, A., & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition*, 3, 227–244.

Ingham, B. (2001). Nominal and verbal status in Lakhota: A lexicographical study. *International journal of American linguistics*, *67*(2), 167-192.

INVALSI (2009). Rilevazione degli Apprendimenti, Questionario Studente, Scuola Primaria, Classe Quinta. Available on https://www.invalsi.it/valsis/docs/Questionario_studente_classe_V_primaria.pdf

ILO International Labour Organization (2008). International Standard Classification of Occupations. Available at https://ilostat.ilo.org/resources/concepts-and-definitions/classification-occupation/

Jaber, A. (2014). *On genericity and definiteness in modern standard Arabic* (Doctoral dissertation). The University of North Carolina at Chapel Hill.

Jaber, A., Al-Momani, R., & Omari, O. (2022). Form-function mismatches: The semantics of the definite article in Jordanian Arabic. *Lingua*, *266*, 103169.

Jakubowicz, C. (2004). Is movement costly? The grammar and the processor in language acquisition. Presentation, JEL (Journée d'Etudes Linguistiques), Nantes.

Jakubowicz, C. (2005). *The language faculty: (Ab)Normal development and interface constraints. Presentation*, GALA 2005, Siena.

Jakubowicz, C., & Strik, N. (2008). Scope-marking strategies in the acquisition of long distance wh-questions in French and Dutch. *Language and Speech*, *51*(1-2), 101-132.

Jenks, P. (2018). Articulated definiteness without articles, *Linguistic Inquiry*, *49*(3), 501-536.

Jia, R., & Paradis, J. (2015). The use of referring expressions in narratives by Mandarin heritage language children and the role of language environment factors in predicting individual differences, *Bilingualism: Language and Cognition*, *18*(4), 737-752.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin speaking children and adolescents: Age-related differences, *Journal of Speech, Language, and Hearing Research,* 50, 1280–1299.

Kail, M., & Hickmann, M. (1992). French children's ability to introduce referents in narratives as a function of mutual knowledge, *First Language*, 12(34), 73–94. DOI: 10.1177/014272379201203405

Kaltsa, M., Tsimpli, I. M., & Argyri, F. (2019). The development of gender assignment and agreement in English-Greek and German-Greek bilingual children. *Linguistic Approaches to Bilingualism*, *9*(2), 253-288.

Kangas, J., & Reunamo, J. (2019). Action Telling Method: From Storytelling to Crafting the Future. In Kerry-Moran, K. J., Aerila, J.-A. (Eds). *Story in Children's Lives: Contributions of the Narrative Mode to Early Childhood Development, Literacy, and Learning, Educating the Young Child* 16, 97-112. https://doi.org/10.1007/978-3-030-19266-2_6

Kantola, L., & van Gompel, R. P. G. (2011). Between-and within-language priming is the same: Evidence for shared bilingual syntactic representations. *Memory & Cognition, 39*(2), 276-290. https://doi.org/10.3758/s13421-010-0016-5

Karmiloff-Smith, A. (1979). A functional approach to child language: A study of determiners and reference. Cambridge: Cambridge University Press

Karmiloff-Smith, A. (1985). Language and cognitive process from a developmental perspective. *Language and Cognitive Processes*, 1, 65–85. DOI: 10.1080/01690968508402071

Karttunen, L. (1968). *What do referential indices refer to*?, v.3854. Rand Corporation. http://www.rand.org/pubs/papers/2008/P3854.pdf.

Karttunen, L. (1969/1976). Discourse referents. In *Proceedings of the 1969 Conference on Computational Linguistics*. S.ng-S.by, Sweden, 1–38. Reprinted in: J. McCawley (ed.). *Syntax and Semantics 7: Notes from the Linguistic Underground.* 363–385. New York: AcademicPress.

Kerry-Moran, K. J., Aerila, J.-A. (Eds) (2019). *Story in Children's Lives: Contributions of the Narrative Mode to Early Childhood Development, Literacy, and Learning, Educating the Young Child* 16, 97-112. https://doi.org/10.1007/978-3-030-19266-2_6

Kašćelan, D., Prevost, P., Serratrice, L., Tuller, L., Unsworth, S., & De Cat, C. (2021). A Review of questionnaires quantifying bilingual experience in children: Do we document the same constructs? *Bilingualism: Language and Cognition*, 25(1), 29-41. https://doi.org/10.1017/S1366728921000390

Kashgary, A. D. (2015). Meanings and Functions of the Arabic Article al-'the'with Reference to the Concept of Definiteness. *Journal of Semitic studies*, *60*(1), 93-110.

Keysar, B., Barr, D., Balin, J., & Bruauner, J. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11, 32–38. DOI: 10.1111/1467-9280.00211

Khan, G. A. (1984). Object markers and agreement pronouns. In *Semitic languages1*. *Bulletin of the School of Oriental and African Studies*, 47(3), 468-500.

Krämer, I. (2003). Reference of indefinite and pronominal noun phrases in a story context: English children's comprehension. In B. Beachley, A. Brown, & F. Conlin (Eds). *Proceedings of BUCLD27*, 449–60. Somerville, MA: Cascadilla Press.

Krifka, M. (1995). Focus and the interpretation of generic sentences. *The generic book*, 238-264.

Kripke, S. (1977). *Speaker's reference and semantic reference*. Midwest Studies in Philosophy 2, 255–276.

Kupisch, T. (2007). Determiners in bilingual German-Italian children: What they tell us about the relation between language influence and language, *Bilingualism: Language and Cognition, 10*, 57–78.

Kupisch, T. & Bernardini, P. (2007). Determiner use in Italian-Swedish and Italian-German children: Do Swedish and German represent the same parameter-setting? In M. Anderssen, & M. Westergaard (Eds). *Papers from the Language Acquisition Workshop, SCL 2006*, 34.3: 209–229. Nordlyd.

Kupisch, T., & Rothman, J. (2016). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers, *International Journal of Bilingualism,* 22, 564–582. doi: 10.1177/1367006916654355

Labov, W. (1982). Objectivity and commitment in linguistic science: The case of the Black English trial in Ann Arbor, *Language in society*, *11*(2), 165-201.

Lambrecht, K. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press

Lautenschläger, T., Schneller, K., Kaiser-Kratzmann, J., & Sachse, S. (2021). *Erfassung der Erzählfähigkeit mehrsprachig aufwachsender Vorschulkinder: Überprüfung der Testgütekriterien des Verfahrens MAIN (Multilingual Assessment Instrument for Narratives)*. Logos, https://doi.org/10.7345/prolog-2103164.

Le Bruyn, B., & Alexandropoulou, S. (2014). Inalienable possession: The semantics of the definite article. In A. Aguilar-Guevara, B. Le Bruyn, & J. Zwarts (Eds). *Weak referentiality*, 311-334. Amsteerdam: John Benjamins.

Leonetti, M. (1999). El articulo. In I. Bosque, & V. Demonte (Eds). *Gramática básica de la lengua española,* 785-890. 1st vol. Madrid: Espasa Clape.

Leonetti, M. (2019). On weak readings of definite DPs. In N. Pomino (Ed). *Proceedings of the IX Nereus International Workshop "Morphosyntactic and semantic aspects of the DP in Romance and beyond*, 1-26. Arbeitspapier 131. Fachbereich Linguistik, Universität Konstanz.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production, *Behavioral and brain sciences*, *22*(1), 1-38.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Levorato, M.C. & Roch, M. (2020). Italian adaptation of the Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics*, 64, 139–146.

Levshina, N. (2015). *How to do linguistics with R*. Amsterdam: John Benjamins Publishing Company.

Li, C.S., & Thompson, S.A. (1981). *Mandarin Chinese: A Functional Reference Grammar*. Berkeley and Los Angeles: University of California Press.

Li W.D. (2014). The pragmatics of existential-presentative constructions in Chinese A discourse-based study. *International Journal of Chinese Linguistics*, 1(2), 244–274.

Liceras, J., Fern.ndez Fuertes, R., & Alba de la Fuente, A. (2012). Subject and copula omission in the English grammar of English-Spanish bilinguals: On the issue of directionality of interlinguistic influence, *First Language*, 32, 88–115.

Lightfoot, D. (1991). *How to set parameters: Arguments from language change* (p. 217). Cambridge, MA: Mit Press.

Lin, Y., & Li, S. (2022). Children's Early Acquisition of Syntactic Category: A Corpus-Based Analysis of English Determiner-Noun Combinational Flexibility. *International Journal of English Linguistics*, *12*(6).

Lindgren, J., Reichardt, V., & Bohnacker, U. (2022). Character introductions in oral narratives of Swedish–German bilingual preschoolers. *First Language*, *42*(2), 234-262.

Lindgren, J., Tselekidou, F., & Gagarina, N. (2023). Acquisition of narrative macrostructure: A comprehensive overview of results from the Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics*, *65*, 111-132.

Listanti, A., Torregrossa, J., Eisenbeiß, S., & Bongartz, C. (2023). Home Literacy Exposure in the Heritage Language Enhances Theory-of-Mind Development: A Study on Greek-Italian Bilingual Children. In P. Gappmayr, & J. Kellog (Eds). *Proceedings of the 47th annuak Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Lleó, C., & Demuth, C. (1999). Prosodic constraints on the emergence of grammatical morphemes: Cross-linguistic evidence from Germanic and Romance languages. In A. Greenhill, H. Littlefield, & C. Tano (Eds). *Proceedings of the 23rd Annual Boston University Conference on Child Language Development*, 407-418. Somerville, MA: Cascadilla Press.

Lleó, C. (1997). Filler syllables, proto-articles and early prosodic constraints in Spanish and German. In A. Sorace, C. Heycock, & R. Shillcock (Eds). *Language acquisition: Knowledge, representation, and processing. Proceedings of the GALA 1997 Conference on Language Acquisition*, 251-256. Edinburgh: University of Edinburgh.

Lleó, C. (2001). The interface of phonology and syntax: The emergence of the article in the early acquisition of Spanish and German. In J. Weissenborn & B. Höhle (Eds), *Approaches to bootstrapping*, 2nd vol., 23-44. Amsterdam: John Benjamins. DOI: 10.1075/lald.24.03lle

Löbner, S. (1985). Definites. *Journal of semantics*, *4*(4), 279-326.

Longobardi, G. (1994). Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic inquiry*, 609-665.

Lonigan, C. J., & Whitehurst, G. J. (1998). Child development and emergent literacy. *Child development*, *69*(3), 848-872.

Luk, G., & Grundy, J. G. (2023). The importance of recognizing social contexts in research on bilingualism. *Bilingualism: Language and Cognition*, *26*(1), 25-27.

Luk, G. (2022). Justice and equity for whom? Reframing research on the "bilingual (dis) advantage". *Applied Psycholinguistics*, *44*(3), 301-315.

Luise, M.C. (2013). Plurilinguismo e multilinguismo in Europa per una Educazione plurilingue e interculturale [Plurilingualism and multilingualism in Europe for plurilingual and intercultural education]*, LEA - Lingue e letterature d'Oriente e d'Occidente.* 2: 525-535.

Luo, J., Yang, W.C., Chan, A., Cheng, K., Kan, R. & Gagarina, N. (2020). The Multilingual Assessment Instrument for Narratives (MAIN): Adding Mandarin to MAIN. *ZAS Papers in Linguistics*, 64, 159–162

Lyons, C. (1999). *Definiteness*. Cambridge University Press, Cambridge.

Lyons, C. (1977). The demonstratives of English, *Belfast Working Papers in Language and Linguistics* 2.5.

Lyons, J. (2002). *Linguistic Semantics: An Introduction*. Cambridge University Press, Cambridge.

Maas, U., & Procházka, S. (2022). Nominal determination in Moroccan Arabic. *Studies in Language*, *46*(4), 793-846.

Macleroy,V. (2022). Siblings' Multilingual Discourse. In Stavans, A., & Jessner-Schmid,, U. (Eds). *The Cambridge Handbook of Childhood Multilingualism.* 325-352. Cambridge: Cambridge University Press.

MacWhinney B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates.

MAIN team at ZAS, (2021), CLAN transcription rules for MAIN [pdf version], available on https://www.leibnizzas.de/fileadmin/media/Dokumente/ZASPiL/ZASPiL_63/MAIN_Transcription_rules.pdf

Malvern, D., Richards, B., Chipere, N., & Purán, P. (2004), *Lexical diversity and language development.* New York: Palgrave Macmillan.

Maratsos, M. P. (1976). The use of definite and indefinite reference in young children: An experimental study of semantic acquisition.

Marçais, P. (1977). Esquisse grammaticale de l'arabe maghrébin. *(No Title)*.

Marini, A., Tavano, A., & Fabbro, F. (2008). Assessment of linguistic abilities in Italian children with specific language impairment. *Neuropsychologia*, *46*(11), 2816-2823.

Martari, Y. (2017). L'acquisizione dell'italiano da parte di sinofoni. Un contributo alla linguistica educativa [The acquisition of Italian by Chinese-speakers. A contribution to educational linguistics.]. Bologna: Pàtron.

Martari, Y. (2021). Basic variety e interlingua in italiano l2. note sulla scrittura di arabofoni [Basic variety and interlanguage in Italian L2. Notes on the writing of Arabic speakers], *Italiano LinguaDue*, *13*(2), 74-95.

Matthews, D., Biney, H., & Abbot-Smith, K. (2018). Individual differences in children's pragmatic ability: A review of associations with formal language, social cognition, and executive functions. *Language Learning and Development*, *14*(3), 186-223.

McCarthy, P. M., & Jarvis, S. (2007). vocd: A theoretical and empirical evaluation. *Language Testing*, *24*(4), 459-488.

Mehotcheva, T. H., & Köpke, B. (2019). Introduction to L2 attrition.

Meir, N., & Novogrodsky, R. (2023). Referential expressions in monolingual and bilingual children with and without Autism Spectrum Disorder (ASD): A study of informativeness and definiteness. *Journal of Child Language*, *50*(2), 215-244.

Meisel, J. M. (1983). Transfer as a second-language strategy, *Language & communication*, *3*(1), 11-46.

Meisel, J. (1986). Word order and case marking in early child language. evidence from simultaneous acquisition of two first languages: French and German, L*inguistics*, *24*(1), 123-184. https://doi.org/10.1515/ling.1986.24.1.123

Meisel, J. (1989). Early differentiation of language in bilingual children. In K. Hyltenstam & L. Obler (Eds), *Bilingualism across a lifespan: Aspects of acquisition, maturity and loss*, 13-40. Cambridge, UK: Cambridge University Press.

Meisel, J. (2001). The simultaneous acquisition of two first languages: Early differentiation and subsequent development of grammars. In J. Cenoz & F. Genesee (Eds), *Trends in bilingual acquisition,* 11-41. Amsterdam: John Benjamins.

Meisel, J. (2007). On autonomous syntactic development in multiple first language acquisition. In H. Caunt-Nulton, S. Kulatilake, & I-H Woo (Eds). *Proceedings of the 31st Boston University Conference in Language Development,* 26-45. Somerville, MA: Cascadilla Press.

Meluzzi, C., Chini, M., & Versino, P. (2018). Due prospettive trasversali [Two transversal perspectives]. In Chini, M., & Andorno, C. (Eds). *Repertori e usi linguistici nell'immigrazione*, 119-160, Milano: Franco Angeli.

Miller, P. J., & Moore, B. B. (1989). Narrative conjunctions of caregiver and child: A comparative perspective on socialization through stories. *Ethos*, *17*(4), 428-449.

Mion, G. (2009). L'indetermination nominale dans les dialectes arabes. Une vue d'ensemble. *Miscellanea arabica*, 215-231.

Mion. G. (2020). L'arabo. Libretto di istruzioni per insegnanti di una classe plurilingue [Arabic. Instruction booklet for teachers of a multilingual class]. In I. Fiorentini, C. Gianollo, & N. Grandi (Eds). *La classe plurilingue*, 175-191. Bologna: Bononia University Press.

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis, *Cognitive Psychology*, 41, 49–100. DOI:10.1006/cogp.1999.0734

Modyanova, N., & Wexler, K. (2007). Semantic and pragmatic language development: Children know 'that'better. In A. Belykova, L. Meroni, & M. Umeda (Eds). *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, 297-308. Somerville, MA: Cascadilla Press.

Moll, L.C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of Knowledge for Teaching: Using a Qualitative Approach to Connect Homes and Classrooms'. *Theory into Practice,* 31(2), 132-141.

Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge University Press.

Moro, F. (2016). *Dynamics of Ambon Malay: Comparing Ambon and the Netherlands*. Unpublished doctoral dissertation. Radboud University, Nijmegen.

Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition, 4*, 1–53.

Napoli, M. (2009). Aspects of Definiteness in Greek. *Studies in Language*, 33(3), 569-611.

Nelson, K. (1985). *Making Sense: The Acquisition of Shared Meaning*. Cambridge (MA): Academic Press.

Nelson, K. (2007). Developing past and future selves for time travel narratives. *Behavioral and Brain Sciences*, *30*(3), 327-328.

Nicoladis, E. (2002). What's the difference between toilet paper and paper toilet? French-English bilingual children's crosslinguistic transfer in compound nouns. *Journal of Child Language, 29,* 843–863.
Nicoladis, E. (2003). Cross-linguistic transfer in deverbal compounds of preschool bilingual children. *Bilingualism: Language and Cognition, 6*, 17–31.

Nicoladis, E. (2006). Cross-linguistic transfer in adjective-noun strings by preschool bilingual children. *Bilingualism: Language and Cognition, 9*, 15–32.

Nicoladis, E. (2012). Crosslinguistic influence in French-English bilingual children's possessive constructions. *Bilingualism: Language and Cognition*, 15(2), 320-328. https://doi.org/10.1017/S1366728911000101

Nicoladis, E., & Gavrila, A. (2015). Crosslinguistic influence in Welsh-English bilingual children's adjectival constructions. *Journal of Child Language*, 42(4), 903-916. https://doi.org/10.1017/S0305000914000440

Nicoladis, E., Rose, A., & Foursha-Stevenson, C. (2010). Thinking for speaking and cross-linguistic transfer in preschool bilingual children. *International Journal of Bilingual Education and Bilingualism*, 13, 345–370.

Nicolopoulou, A., Cortina, K. S., Ilgaz, H., Cates, C. B., & de Sá, A. B. (2015). Using a narrative-and play-based activity to promote low-income preschoolers' oral language, emergent literacy, and social competence. *Early childhood research quarterly*, *31*, 147-162.

Nicolopoulou, A. (2019). Using a storytelling/story-acting practice to promote narrative and other decontextualized language skills in disadvantaged children. *Narrative, literacy and other skills: Studies in intervention*, 263-284.

Ninio, A., & Snow, C. E. (1999). The development of pragmatics: Learning to use language appropriately. In *Handbook of child language acquisition*, 347-383. Brill.

Otwinowska, A., Mieszkowska, K., Białecka-Pikul, M., Opacki, M., and Haman, E. (2020). Retelling a model story improves the narratives of polish-English bilingual children. *International Journal of Bilingual Education and Bilingualism*. 23, 1083–1107. doi: 10.1080/13670050.2018.1434124

Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism*, 5(1), 19-38. https://doi.org/10.1177/13670069010050010201

Paradis, J. (2011). Internal and external factors in child second language acquisition. *Linguistic Approaches to Bilingualism*. 1, 213–237. DOI: 10.1075/lab.1.3.01par

Paradis, J. (2017). Parent report data on input and experience reliably predict bilingual development and this is not trivial. *Bilingualism: Language and Cognition*, 20 (1), 2017, 27–28, Cambridge University Press. 2016 DOI:10.1017/S136672891600033X.

Paradis, J. (2023). Sources of individual differences in the dual language development of heritage bilinguals. *Journal of Child Language*, *50*(4), 793-817.

Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition, 18*, 1–25.

Paradis, J., & Navarro, S. (2003). Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of the input? *Journal of Child Language*, 30(2), 371-393. https://doi.org/10.1017/S0305000903005609

Paris, M. C. (2007). Un aperçu de la réduplication nominale et verbale en mandarin. *Faits de langues*, *29*(1), 63-76.

Partee, B. H. (1970). Opacity, coreference, and pronouns. *Synthese*, 21, 359-385.

Patuto, M., Repetto, V., & Müller, N. (2011). Delay and acceleration in bilingual first language acquisition: The same or different? In E. Rinke & T. Kupisch (Eds), *The development of grammar, language acquisition and diachronic change* , 231-259. Amsterdam: John Benjamins.

Paul R., Hernandez R., Taylor L., & Johnson K. (1996). Narrative development in late talkers: Early school age. *Journal of Speech and Hearing Research*, 39, 1295-1303.

Payant, C., & Galante, A. (2022). A Word From The Guest Editors: Plurilingualism and Translanguaging: Pedagogical Approaches for Empowerment and Validation—An Introduction. *TESL Canada Journal*, *38*(2), i-xxii.

Pérez-Leroux, A., Pettibone, E., & Castilla-Earls, A. (2017). Down two steps: Are bilinguals delayed in the acquisition of recursively embedded PPs?. *Matraga-Revista do Programa de Pós-Graduação em Letras da UERJ*, *24*(41), 393-416.

Pérez-Leroux, A. T., & Battersby, T. (2009). Early determination. *Hispanic child languages*, 117-139.
Pérez-Leroux, A.T., Cuza, A., & Thomas, D. (2011). Clitic placement in Spanish/English bilingual children. *Bilingualism: Language and Cognition*, 14, 221–232.

Pérez -Leroux, A.T., Pirvulescu, M., & Roberge, Y. (2009). Bilingualism as a window into the language faculty: The acquisition of objects in French-speaking children in bilingual and monolingual contexts. *Bilingualism: Language and Cognition*, 12, 97–112.

Pesonen, J. (2019). Children's Stories Supporting the Development of Critical Literacy and Intercultural Understanding. *Story in Children's Lives: Contributions of the Narrative Mode to Early Childhood Development, Literacy, and Learning*, 275-297.

Piccardo, E., Germain-Rutherford, A., & Lawrence, G. (2021). *The Routledge handbook of plurilingual education*. Routledge.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*(4), 633-651. https://doi.org/10.1006/jmla.1998.2592

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, *27*(2), 169-190.

Pine, J. M., & Lieven, E. V. M. (1997). Slot and frames patterns and the development of the determiner category. *Applied Psycholinguistics*, 18(2), 123–138. DOI: 10.1017/S0142716400009930

Pine, J. M., Freudenthal, D., Krajewski, G., & Gobet, F. (2013). Do young children have adult-like syntactic categories? Zipf's law and the case of the determiner. *Cognition*, 127(3), 345–360. DOI: 10.1016/j.cognition.2013.02.006

Pirvulescu, M., Pérez-Leroux, A. & Roberge, Y. (2012). A bidirectional study of object omissions in French-English bilinguals. In K. Braunmüller, C. Gabriel, & B. H.nel-Faulhaber

(Eds), *Multilingual individuals and multilingual societies*, 171–188. Amsterdam: John Benjamins.

Pirvulescu, M., Pérez-Leroux, A. T., Roberge, Y., Strik, N., & Thomas, D. (2014). Bilingual effects: Exploring object omission in pronominal languages. *Bilingualism: Language and Cognition*, *17*(3), 495-510.

Pizzuto, E., & Caselli, M. C. (1992). The acquisition of Italian morphology: Implications for models of language development. *Journal of child language*, *19*(3), 491-557.

Prince, E. (1981). On the inferencing of indefinite-this NPs. In Joshi, A., Weber, B.H., & Sag, I.A., (eds.). *Elements of Discourse Understanding*. Cambridge University Press. 231-250.

Polinsky, M. (2006). Incomplete acquisition: American Russian. J. *Slavic Linguist*. 14, 191–262.

Puroila, A. M. (2019). Who am I? Shaping young children's identities through everyday narratives. *Story in children's lives: Contributions of the narrative mode to early childhood development, literacy, and learning*, 55-75.

Pugliese, R. (2015). «Tradurre per la compagna di banco: child language brokering e interazioni costruttive nella classe plurilingue». Corrà, L. (a cura di), Educazione linguistica in classi multietniche = Atti GISCEL 2014-2015. Roma: Aracne, 85-116.

Quay, S., & Montanari, S. (2016). Early bilingualism: From differentiation to the impact of family language practices. In E. Nicoladis & S. Montanari (Eds), *Bilingualism across the lifespan: Factors moderating language proficiency*, 23-42. American Psychological Association. https://doi.org/10.1037/14939-003

R Core Team. (2021). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. https://www.R-project.org/

Radford, A. (1990). The syntax of nominal arguments in early child English. *Language Acquisition*, 1(3), 195–223. DOI: 10.1207/s15327817la0103_1;

Renzi, L., (1988). L'articolo [The Article]. In L. Renzi, G. Salvi & A. Cardinaletti (Eds). *Grande grammatica italiana di consultazione*, 1st vol., 371-437. Bologna: il Mulino.

Reichardt V. (2014). *Wie führen Kinder Referenten ein? Eine Untersuchung der Verwendung deutscher Indefinita durch russisch-deutsch bilinguale Kinder verschiedener Altersgruppen* [How do children introduce referents? A study of the use of German indefinites by Russian–German bilingual children in different age groups] [Unpublished BA thesis]. Humboldt-Universität zu Berlin, Berlin.

Restrepo, M.A., Castilla, A. P., Schwanenflugel, P. J., Neuharth-Pritchett, S., Hamilton, C. E., & Arboleda, A. (2010). Effects of a supplemental Spanish oral language program on sentence length, complexity, and grammaticality in Spanish-speaking children attending English-only preschools. *Language, Speech and Hearing Services in Schools*, 41(1), 3–13. https://doi.org/10.1044/2018_JSLHR-L-18-0324

Roberts, I. (2001). Language change and learnability. *Language acquisition and learnability*, *81*, 125.

Roch, M., Florit, E., & Levorato, C. (2012). Questionario per genitori di bambini bilingui. Available at https://www.dpss.unipd.it/lab-las/materials/questionnaires.

Roeper, T. (2006). Watching noun phrases emerge. Seeking compositionality. In V. van Geenhoven (Ed). *Acquisition meets semantics*. Dordrecht: Kluwer Academic Publishers.

Rolstad, K., & MacSwan, J. (2014). The facilitation effect and language thresholds. *Frontiers in psychology*, *5*, 1197.

Rowlands, E. C. (1969). *Yoruba*. London: Hodder and Stoughton.

Rozendaal. M.L. (2008). *The acquisition of reference: A cross-linguistic study*. Netherlands Graduate School of Linguistics.

Rozendaal, M. I., & Baker, A. E. (2008). A cross-linguistic investigation of the acquisition of the pragmatics of indefinite and definite reference in two-year-olds. *Journal of Child Language,* 35(4), 773–807. DOI: 10.1017/S0305000908008702

Rubio-Fernandez, P. (2021). Pragmatic markers: the missing link between language and Theory of Mind. *Synthese*, *199*(1-2), 1125-1158.

Runnqvist, E., Gollan, T. H., Costa, A., & Ferreira, V. S. (2013). A disadvantage in bilingual sentence production modulated by syntactic frequency and similarity across languages. *Cognition*, *129*(2), 256-263.

Russell, B. (1905). On denoting. *Mind* 14: 479–93.

Saccardo, D. (2016). La politica linguistica nella scuola italiana [Language policy in Italian schools]. In C.A. Meleros Rodrìguez (Ed). *Le lingue in Italia, le lingue in Europa: dove siamo, dove andiamo,* 21-29*,* Venezia, Edizioni Ca' Foscari - Digital Publishing Università Ca' Foscari Venezia.

Salem, M. (2010). Bare nominals, information structure and word order, *Lingua*, 120(6), 1476-1501.,

Schaeffer, J., & Matthewson, L. (2005). Grammar and pragmatics in the acquisition of article systems, *Natural Language and Linguistic Theory*, 23, 53–101. DOI: 10.1007/s11049-004-5540-1

Schafer, R., & de Villiers, J. (2000). Imagining articles: What a and the can tell us about the emergence of DP. In S. C. Howell, S. Fish, & T. Keith-Lucas (Eds). *Proceedings of BUCLD 24*, 609-620. Sommerville, MA: Cascadilla Press

Scheele, A. F., Leseman, P. P., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, *31*(1), 117-140.

Schimke, S., Colonna, S. and Hickmann, M. (2015). The development of discourse cohesion in French and German: General and language-specific determinants. *Child Language Symposium*. University of Warwick, Coventry 20-21 Juli 2015.

Schimke, S., Colonna, S. and Hickmann, M. (2020). Reference in French and German. In Gagarina, N., Musan, R. (Eds*), Referential and Relational Discourse Coherence in Adults and Children*. 139-170. Mouton De Gruyter.

Schmitt, B. M., Meyer, A. S., & Levelt, W. J. (1999). Lexical access in the production of pronouns. *Cognition*, *69*(3), 313-335.

Schulz, P. & Grimm, A. (2019) The Age Factor Revisited: Timing in Acquisition Interacts With Age of Onset in Bilingual Acquisition. *Frontiers in Psychology*. 9: 2732.doi: 10.3389/fpsyg.2018.02732

Schwartz, M., & Rovner, H. (2015). The acquisition of definiteness in Hebrew (L2) by bilingual preschool children (Russian-L1): A longitudinal multiple-case study. *International Journal of Bilingualism*, *19*(5), 548-571.

Schwarz, F. (2013). Two kinds of definites cross-linguistically. *Language and Linguistics Compass*, 7(10), 534–559.

Selkirk, E.O. (1984). *Phonology and syntax: the relation between sound and structure*. Cambridge, Mass.: MIT Press.

Serratrice, L. (2006). The role of perceptual and discourse cues in the choice of referential expressions in English pre-schoolers and school-age children. *Leeds Working Papers in Linguistics & Phonetics*, 11, 173–82.

Serratrice, L. (2007). Cross-linguistic influence in the interpretation of anaphoric and cataphoric pronouns in English–Italian bilingual children. *Bilingualism: Language and Cognition*, 10, 225–238.

Serratrice L. (2009). Carving up referential space: A priming account of pronoun use in Italian. In J. Chandelee, M. Franchini, S. Lord, & G.M. Rheiner (Eds). *Proceedings of the 33rd Annual Boston University Conference on Language Development*, 468-479. Sommerville, MA: Cascadilla Press

Serratrice, L. (2013). Crosslinguistic influence in bilingual development: Determinants and mechanisms, *Linguistic Approaches to Bilingualism*, 3(1), 3-25. https://doi.org/10.1075/lab.3.1.01ser

Serratrice, L. (2016). Crosslinguistic influence, crosslinguistic priming and the nature of shared syntactic structures, *Linguistic Approaches to Bilingualism*, 6(6), 822-827. https://doi.org/10.1075/lab.6.6.15ser

Serratrice, L. (2022). What can syntactic priming tell us about crosslinguistic influence?. In K. Messenger (Ed). *Syntactic priming in language acquisition: representations, mechanism and applications*, 129-156. Amsterdam: John Benjamins, 156. ISBN:9789027211637 | doi: https://dx.doi.org/10.1075/tilar.31.07ser

Serratrice, L., & De Cat, C. (2020). Individual differences in the production of referential expressions: The effect of language proficiency, language exposure and executive function in bilingual and monolingual children, *Bilingualism: Language and Cognition*, *23*(2), 371-386.

Serratrice, L., Sorace, A., Filiaci, F., & Baldo, M. (2009). Bilingual children's sensitivity to specificity and genericity: Evidence from metalinguistic awareness, *Bilingualism: Language and Cognition*, 12, 239–257.

Serratrice, L., Sorace, A., Filiaci, F., & Baldo, M. (2012). Pronominal objects in English–Italian and Spanish–Italian bilingual children, *Applied Psycholinguistics*, 33(4), 725–751. https://doi.org/10.1017/S0142716411000543

Serratrice, L., Sorace, A., & Paoli, S. (2004). Crosslinguistic influence at the syntax-pragmatics interface: Subjects and objects in English-Italian bilingual and monolingual acquisition, *Bilingualism: Language and Cognition*, 7, 183–205..

Silva- Corvalàn, C., & Montanari, S. (2008). The acquisition of ser, estar (and be) by a Spanish-English bilingual child: The early stages. *Bilingualism: Language and Cognition,* 11, 341–360.

Skarabela, B., Allen, S.E.M., Scott-Phillips, T.C. (2013). Joint attention helps explain why children omit new referents. *Journal of Pragmatics*, 56, 58–69. DOI: https://doi.org/10.1016/j.pragma.2012.08.003

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22, 339–368.

Sorace, A., Filiaci, F., Serratrice, L., & Baldo, M. (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, *119*, 460–477.

Sorace, A., Serratrice, L., Filiaci, F., & Baldo, M. (2009). Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua, 119*(3), 460-477. https://doi.org/10.1016/j.lingua.2008.09.008

Sparvoli, C. (2017). Fenomeni di inversione in cinese e in italiano [Inversion phenomena in Chinese and Italian]. In C. Bulfoni (Ed). *Linguistica Cinese. Tendenze e Prospettive*, 113-142. Milan: Unicopli.

Stadler M.A. & Ward G.C. (2005). Supporting the Narrative Development of Young Children, *Early Childhood Education Journal*, 33(2), 73-80.

Stein N.L., Glenn C.G. (1979). An analysis of story comprehension in elementary school children. In Freedle, Roy Omer (Ed). *New directions in discourse processing*. Hillsdale, NJ: Ablex Publishing Corporation.

Strik, N. (2009). Derivational contrasts in Dutch and French. *Linguistics in the Netherlands*, *26*, 91-102.

Strik, N., & Pérez -Leroux, V. (2011). Jij doe wat girafe? Wh-movement and inversion in Dutch- French bilingual children. *Linguistics Approaches to Bilingualism*, 1, 175–205.

Sun, H., & Yin, B. (2020). Multimedia input and bilingual children's language learning. *Frontiers in Psychology*, *11*, 2023.

Surrain, S., & Luk, G. (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, *22*(2), 401-415.

Swain, M. (1972). *Bilingualism as a first language*. Unpublished doctoral dissertation, University of California, Irvine.

Tartar, M. (2009). *Enchanted hunters: The power of stories in childhood*. New York: W. W. Norton.

Titone, D. A., & Tiv, M. (2023). Rethinking multilingual experience through a Systems Framework of Bilingualism. *Bilingualism: Language and Cognition*, *26*(1), 1-16.

Tolchinsky, L. (2008). The Emergence of Writing. In MacArthur, C.A., Graham, S.& Fitzgerald, J. (Eds). *Handbook of Writing Research*. 293-310. Guilford Press.

Tomasello M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Harvard University Press.

Tomasello, M. (2018). How children come to understand false beliefs: A shared intentionality account. In N. Raikhel (Ed). *Proceedings of the National Academy of Sciences*, 115(34), 8491-8498. Washington DC: NAS.

Tomić A, Rodina Y, Bayram F & DeCat,C. (2023). Documenting heritage language experience using questionnaires. *Frontiers in Psychology*. 14:113-1374. DOI: 10.3389/fpsyg.2023.1131374

Topaj N. (2010). Topical referential expressions in narratives of Russian-German bilingual children. In Chini M. (Ed). *Topic, information structure and language acquisition*, 59-72. Milan: Franco Angeli.

Torregrossa, J. (2017). The role of executive functions in the acquisition of reference: The production of demonstrative pronouns by German monolingual children. In J. Choi, H. Demirdache, O. Lungu & L. Voeltzel (Eds). *Language acquisition at the interfaces: Proceedings of GALA 2015*, 316-330. New Castle upon Tyne: Cambridge Scholars Publishing.

Torregrossa J, Andreou M, Bongartz C, Tsimpli I.M. (2021). Bilingual acquisition of reference: The role of language experience, executive functions and crosslinguistic effects, *Bilingualism: Language and Cognition*, 24, 694–706. https://doi.org/10.1017/ S1366728920000826

Torregrossa J., Bongartz C. (2018). Teasing Apart the Effects of Dominance, Transfer, and Processing. Reference Production by German–Italian Bilingual Adolescents, *Languages*, 3, 36. DOI:10.3390/languages3030036

Torregrossa J., Carbonara V. (2022). Le scelte linguistiche di alcune famiglie con background migratorio in Italia: Usi, competenze e atteggiamenti nei confronti del multilinguismo [The linguistic choices of some families with a migratory background in Italy: Uses, skills and attitudes towards multilingualism]. *Studi d'italianistica nell'Africa australe Italian studies in Southern Africa*, 35(2), 87-120.

Towson, J. A., Fettig, A., Fleury, V. P., & Abarca, D. L. (2017). Dialogic reading in early childhood settings: A summary of the evidence base. *Topics in Early Childhood Special Education*, *37*(3), 132-146.

Treffers-Daller, J., Özsoy, A. S., & Van Hout, R. (2007). (In) complete acquisition of Turkish among Turkish–German bilinguals in Germany and Turkey: An analysis of complex embeddings in narratives. *International Journal of Bilingual Education and Bilingualism*, *10*(3), 248-276.;

Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English.,*International Journal of Bilingualism, 8*, 257–277.

Tsimpli, I. M. (2014). Early, late or very late?. *Linguistic Approaches to Bilingualism*, 4, 283–313. doi: 10.1075/lab.4.3.01tsi

Turner, M. (2021). Definiteness systems and dialect classification. *Languages*, *6*(3), 128.

UNESCO. (2011). International Standard Classification of Education. Available at ISCED-11 available on http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf

Unsworth, S. (2012). Testing for crosslinguistic influence and exposure effects in the bilingual acquisition of specific indefinite objects. In A. K. Biller, E. Y. Chung, & A. E. Kimball (Eds). *Proceedings of the 36th annual Boston University Conference on Language Development (BUCLD 36)*, 633-645. Sommerville, MA: Cascadilla Press.

Unsworth, S. (2016). Quantity and Quality of Language Input in Bilingual Language Development. In Nicoladis, E., & Montanari, S. (Eds). *Bilingualism Across the Lifespan: Factors Moderating Language Proficiency,* American Psychological Association and Walter de Gruyter, GmbH.

Unsworth, S. (2019). Quantifying experience in heritage language development. In M. Schmid & B. Köpke (Eds). *The Oxford handbook of first language attrition*. Oxford: Oxford University Press.

Unsworth, S. (2023). Shared syntax and cross-linguistic influence in bilingual children: Evidence from between- and within-language priming. *Linguistic Approaches to Bilingualism.* https://doi.org/10.1075/lab.22093.uns Also available to read at: 10.31219/osf.io/a3nc9

Usai M.C., Traverso L., Gandolfi E., Viterbori P. (2017). FE-*PS 2-6 Batteria per la valutazione delle funzioni esecutive in età prescolare*, Trento: Erickson.

Valian, V. (1986). Syntactic categories in the speech of young children, *Developmental Psychology*, 22(4), 562–579. DOI: 10.1037/0012-1649.22.4.562.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children, *Cognition*, 40(01), 21–81. DOI: 10.1016/0010-0277(91)90046-7

Valian, V. (2013). Determiners: An empirical argument for innateness. In Sanz, M., Laka, I., & Tanenhaus, M. (Eds). Language down the garden path: *The cognitive and biological basis for*

linguistic structure (Chapter 14). New York: Oxford University Press.

Valian, V. (2015). Bilingualism and cognition, *Bilingualism: language and cognition*, *18*(1), 3-24.

Valian, V. (2020). Variability: Definitions of language and language learning. *Bilingualism: Language and Cognition*, *23*(1), 48-49.

Valian, V., Solt, S., & Stewart, J. (2009). Abstract categories or limited-scope formulae? The case of children's determiners. *Journal of Child Language*, 36(4), 743–778. DOI: 10.1017/S0305000908009082

Van Dijk, C., Van Wonderen, E., Koutamanis, E., Kootstra, G. J., Dijkstra, T., & Unsworth, S. (2021). Cross-linguistic influence in simultaneous and early sequential bilingual children: a meta-analysis. *Journal of Child Language*, 1-33. https://doi.org/10.1017/S0305000921000337

Vasilyeva, M., Waterfall, H., G.mez, P., G.mez, L., Bowers, E., & Shimpi, P. (2010). Crosslinguistic syntactic priming in bilingual children. *Journal of Child Language*, 37, 1047–1064.

Veneziano, E., & Nicolopoulou, A. (Eds). (2019). *Narrative, literacy and other skills: studies in intervention* (Vol. 25). John Benjamins Publishing Company.

Vergnaud, J. R., & Zubizarreta, M. L. (1992). The definite determiner and the inalienable constructions in French and in English. *Linguistic inquiry*, *23*(4), 595-652.

Vihman, M. (1985). Language differentiation by the bilingual infant. Journal of Child Language, 12, 297-324.

Vihman, M., & McLaughlin, B. (1982). Bilingualism and second language acquisition in preschool children. In C. Brainerd & M. Pressley (Eds). *Verbal processes in children*, 35-58. New York: Springer-Verlag.

Volterra, V., & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language*, 5(2), 311-326. https://doi.org/10.1017/S0305000900007492

Von Heusinger, K. (2011). Specificity, referentiality and discourse prominence: German indefinite demonstratives. In *Proceedings of Sinn und Bedeutung*. 15, 9-30.

Von Heusinger, K. (2019). 3. Specificity. In P. Portner, K. Heusinger & C. Maienborn (Ed.), *Semantics - Noun Phrases and Verb Phrases*, 70-111. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110589443-003

Wang, Q., & Leichtman, M. D. (2000). Same beginnings, different stories: A comparison of American and Chinese children's narratives. *Child development*, *71*(5), 1329-1346.

Wellman, R.L., Lewis, A.B., Freebairn, L.A., Avrich, A.A., Hansen A.J. & Stein C.M. (2011). Narrative Ability of Children With Speech Sound Disorders and the Prediction of Later Literacy Skills. Language Speech, and Hearing Services in Schools. 42 (4), 561-579. https://doi.org/10.1044/0161-1461(2011/10-0038)

Westby, C. (1984). Development of narrative language abilities. In Wallach, G., & Butler, K. (Eds). *Language learning disabilities in school- age children*. Baltimore, MD: Williams & Wilkins, 103–127.

Westby C. E., (1991). Learning to talk—talking to learn: Oral-literate language differences In C. S. Simon, (Eds). *Communication skills and classroom success*, 334–357. Eau Claire, WI Thinking Publications.

Wexler, K. (2011). Cues don't explain learning: Maximal trouble in the determiner system. *The processing and acquisition of reference*, *15*.

Whitely, C., & Colozzo, P. (2013). Who's who? Memory updating and character reference in children's narratives.

Wildenger, L. K., McIntyre, L. L., Fiese, B. H., & Eckert, T. L. (2008). Children's daily routines during kindergarten transition. *Early Childhood Education Journal*, *36*, 69-74.

Whitehurst, G. J., & Lonigan, C. J. (2001). Emergent literacy: Development from prereaders to readers. *Handbook of early literacy research*, *1*, 11-29.

Wolfram, W. (1993). Ethical Considerations in Language Awareness Programs.

Wolleb, A., Sorace, A., & Westergaard, M. (2018). Exploring the role of cognitive control in syntactic processing: Evidence from cross-language priming in bilingual children. *Linguistic Approaches to Bilingualism, 8*(5), 606-636. https://doi.org/10.1075/lab.17002.wol

Wright, S. E., & Givón, T. (1987). The pragmatics of indefinite reference: Quantified text-based studies. *Studies in Language*. International Journal sponsored by the Foundation "Foundations of Language", 11(1), 1-33.

Yip, V., & Matthews, S. (2000). Syntactic transfer in a Cantonese-English bilingual child. *Bilingualism: Language and Cognition*, 3, 193–208.

Yule, G. (2013). *Referential communication tasks*. Routledge.

Yurtsever, A., Anderson, J. A., & Grundy, J. G. (2023). Bilingual children outperform monolingual children on executive function tasks far more often than chance: An updated quantitative analysis. *Developmental Review*, *69*, 101084.

Zamparelli, R. (2008). Dei ex machina: a note on plural/mass indefinite determiners. *Studia Linguistica*, 62(3), 301-327.

Zdorenko, T., & Paradis, J. (2012). Articles in child L2 English: When L1 and L2 acquisition meet at the interface. *First language*, *32*(1-2), 38-62.

Zhou, J., Mai, Z., Cai, Q., Liang, Y., & Yip, V. (2022). Reference production in Mandarin–English bilingual preschoolers: Linguistic, input, and cognitive factors. *Frontiers in Psychology,* 13, 897031.

APPENDIX A

Tab.1 Italian article system

	DEFINITE		INDEFINITE	
	SINGULAR	PLURAL	SINGULAR	PLURAL
MASCULINE	*il; lo; l'*	*i; gli*	*un; uno*	
FEMININE	*la; l'*	*le*	*una; un'*	

Tab.2 Italian system of partitive articles

	SINGULAR	PLURAL
MASCULINE	*del; dello; dell'*	*dei; degli*
FEMININE	*della; dell'*	*delle*

Tab.3 Spanish article system

	DEFINITE		INDEFINITE	
	SINGULAR	PLURAL	SINGULAR	PLURAL
MASCULINE	*el (lo)*	*los*	*un*	*unos*
FEMININE	*la; el*	*las*	*una*	*unas*

APPENDIX B

Arabic adaptations of the script of the “Baby Birds” story

Egyptian dialect

اسم القصة: القطة

1,2) الصورة الأولى و الثانية:
دا بيوم من الأيام كان في قطة رخمة بتبص على فراشة صفرا واقفه على ورده .نطت لقدام عايزه تمسكها .بالوقت دا كان فيه ولد فرحان

راجع من صيد السمك و في إيده جردل و كورة. ببص القطة بتجري ورا الفراشة .

3,4)الصورة التالتة و الرابعة: طارت الفراشة بسرعة فوقعت القطة على الوردة تعورت و تضايقت أوي. اتفزع الولد أوي فوقعت الكورة من أيده و وقت بص و شاف الكورة بتتدحرج ناحية الميه زقع قال" يا لهوي ضاعت الكورة بتاعتي " زعل و كان عاوز يرجع كورته . في الوقت دا لاحظت القطة جردل الولد و فكرت "عايزه أخذ سمكة"

5,6)في الصورة الخامسة و السادسة : دا الوقتي كان الواد بيسحب كورته من الميه بالسنارة بتاع الصيد و ماخدش باله أن القطة خذت سمكة في الأخر فرحت القطة لأنها أكلت سمكةً زاكية، و الواد اتبسط أوي عشان رجعت كورته.

Standard

اسم القصة: الهرة

1,2) الصورة الأولى و الثانية:
في يوم من الأيام كان هناك قطة مشاكسة، رأت فراشة صفراء تقف على زهرة، فقفزت إلى الأمام لكي تمسك بها. وفي نفس الوقت كان هناك صبي يشعر بالسعادة لأنه عائد من صيد السمك و يحمل بيديه دلو و كرة ، رأى القطة تلاحق الفراشة.

3,4) في الصورة الثالثة و الرابعة :
طارت الفراشة بسرعة فسقطت الهرة على الزهرة تألمت و استأت جداً. تفاجئ الصبي! فسقطت الكرة من يده و شاهدها تتدحرج نحو الماء صرخ قائلاً " يا إلهي... لقد ضاعت الكرة " حزن و أراد أن يستعيد كرته . في ذات الوقت لاحظت القطة دلو الصبي و فكرت "أريد أن آكل السمكة".

5,6) في الصورة الخامسة و السادسة :عندما بدأ الصبي في سحب كرته من الماء بسنارة الصيد لم ينتبه إلى أن القطة أخذت السمكة ،في النهاية فرحت القطة كثيراً لأنها أكلت سمكة شهية و فرح الصبي لأنه استعاد كرته .

Moroccan dialect

اسم القصة المشة

1,2) الصورة الاولى و الثانية: في واحد من النهار كانت واحد القطيطة قبيحة شافت فراشه صفرا قاعدة على وريدة ،و هي تنقز لقدام حيث تبغى تشدها في نفس الوقت كان واحد الوليد فرحان راجع من الصيد هز ايديه سطل و كورة شاف القطيطة باركة تتبع الفراشة

3,4) في الصورة الثالثة و الرابعة: طارت الفراشة و المشيشة طاحت على هذيك الوريدة تقلقت و توعدت بالزاف و من يلي شافها دريت تخلع عليها ،و هي تطيح ليه الكورة من أيده و من يلي شاف الكورة غات لجهة الماء و هو يغوت: "يااي مشت كورتي" قلق و بغى كورته بهذا الوقت القطيطة شافت السطل ديال الدري و هي تفكر : "أنا نبغى نصيد حوته"

5,6) في الصورة الخامسة و السادسة : الوليد بدا كيجرف الكورة ديالو من الماء بالسنارة ديال الصيد و ماردش البال بإلا القطيطة شدت حوته و القطيطة فرحت بالزاف حيت أكلت واحد الحوته بنينه و الوليد فرح فرح بالزف لأنه رجعت ليه كورته.

Syrian dialect

اسم القصة : القطة

1,2) الصورة الأولى و الثانية :بيوم من الأيام كان في بسة مشاغبة شافت فراشة صفرا واقفة على ورده، فنطت لقدام لأن بدها تمسكها .وبهل الوقت كان في ولد مبسوط راجع من صيد السمك و بيديه دلو و طابة ،صار يراقب القطة و هيي عم تلاحق الفراشة

3,4) في الصورة التالتة و الرابعة: طارت الفراشة بسرعة، فوقعت البسة على الورده توجعت و تضايقت كتير. نقز الصبي كتير !

فوقعت الطابة من إيده صار يشوفها و هي عم تكرج قدامه رايحة للمي ،صرخ و صار يقول "ييبي راحت الطابة " .زعل و صار بده طابته ترجع .بهاد الوقت انتبهت البسة لدلو الولد و فكرت:" بدي أكل سمكة".

5,6) في الصورة الخامسة و السادسة: بلش الصبي يسحب طابته من المي بسنارة الصيد و ما انتبه أنو البسة أخذت السمكة ،بالأخير البسة فرحت كتير لأنو أكلت سمكة طيبة و الولد انبسط كتير كتير لانو رجعتلو طابته.

Questionnaire for the family

QUESTIONARIO PER LA FAMIGLIA

A. [DATI ANAGRAFICI E FAMILIARI]

1. Compila le tabelle con i dati richiesti

	BAMBINO/A
Nome e Cognome	
Sesso	☐ **M** ☐ **F**
Data di nascita	
Luogo di nascita	
Se nati all'estero:	
Data di arrivo in Italia	
Luogo di arrivo in Italia	

	Madre	Padre
Data di nascita		
Luogo di nascita		
Data di arrivo in Italia		
Livello di istruzione	☐ **nessuno** ☐ **educazione primaria** ☐ **educazione secondaria** ☐ **università**	☐ **nessuno** ☐ **educazione primaria** ☐ **educazione secondaria** ☐ **università**
Lavoro		

2. In quale città abita ora il bambino/la bambina? _____
3. Il bambino/la bambina ha vissuto in altre città prima di Bologna? ☐ SI ☐ NO
4. Se SÌ, quale/i? _____
5. Quante persone abitano con il bambino/la bambina? _____
6. Quali persone abitano con il bambino/la bambina: *(puoi mettere più di una x)*

☐ madre ☐ fratello *indica il numero_____* ☐ nonno ☐ altro *(specifica se vuoi)* _____

☐ padre ☐ sorella *indica il numero_____* ☐ nonna

7. Se uno dei genitori non vive con il bambino/la bambina, dove vive? _____
8. Se uno o più fratelli/sorelle non vive con il bambino/la bambina, dove vive? _____
9. Indicare quanti anni hanno i fratelli e/o le sorelle:

frat. / sor. 1: _____ frat. / sor. 2: _____ frat. / sor. 3: _____ frat. / sor.4: _____

B. [SCOLARIZZAZIONE]

1. Il bambino/la bambina è andato/a all'asilo nido? ☐ SI ☐ NO
2. Se Sì, dove? _____
3. è andato/a alla scuola d'infanzia? ☐ SI ☐ NO
4. Se Sì, dove? _____
5. Se sì, per quanto tempo?

☐ alcuni mesi ☐ un anno ☐ due anni ☐ tre anni

(specificare numero _____)

C. [ESPOSIZIONE LINGUISTICA]

1. Che lingue usa e sente il bambino/la bambina?

	il bambino, che lingua parla con	**il bambino, che lingua sente da…**
la MADRE		
il PADRE		
FRATELLO/SORELLA		
AMICI DELLA STESSA COMUNITÀ D'ORIGINE		
AMICI ITALIANI		
MAESTRE		
CUGINI		
ALTRE PERSONE DELLA FAMIGLIA *(specificare)*		

2. Che lingua sente più spesso il bambino/la bambina?

Quando guarda la TV	☐ lingua d'origine	☐ italiano
Quando gioca con gli amici a casa	☐ lingua d'origine	☐ italiano
Quando gioca con gli amici fuori	☐ lingua d'origine	☐ italiano
Quando sta a scuola	☐ lingua d'origine	☐ italiano
Quando sta fuori con i genitori	☐ lingua d'origine	☐ italiano
Quando guarda video sul tablet/telefono/computer	☐ lingua d'origine	☐ italiano

3. Che lingua parla il padre: **a casa** _____ **fuori**: _____
4. Che lingua parla la madre: **a casa** _____ **fuori**:_____

5. A che età il bambino ha sentito l'italiano per la prima volta?

□ dalla nascita □ a _____ mesi *(specificare)* □ a 1 anno □ a 2 anni □ a 3 anni
□ a 4 anni □ a 5 anni □ a 6 anni

6. A che età il bambino ha sentito la lingua d'origine per la prima volta?

□ dalla nascita □ a _____ mesi *(specificare)* □ a 1 anno □ a 2 anni □ a 3 anni
□ a 4 anni □ a 5 anni □ a 6 anni

7. Quale lingua ha parlato per prima? □ lingua d'origine □ italiano
8. Ogni quanto tornate nel paese d'origine?

□ una volta all'anno □ raramente □ mai □ più volte all'anno

9. Che lingua parla il bambino con le persone del paese d'origine?

□ lingua d'origine □ italiano

10. Quanto conosce l'italiano … *(metti una x sotto il numero che descrive meglio il tuo livello)*

	MADRE				PADRE			
	0 per niente	1 poco	2 bene	3 molto bene	0 per niente	1 poco	2 bene	3 molto bene
PARLARE								
CAPIRE chi parla								
CAPIRE lo scritto								
SCRIVERE								

D. **[MATERIALE E ATTIVITÀ DIDATTICI IN CASA]**

1. Avete un pc o un tablet? □ SI □ NO
2. Avete smartphone? □ SI □ NO
3. Quando il bambino/la bambina gioca da solo/a che lingua parla?

□ lingua d'origine □ italiano

4. Con che frequenza svolgete queste attività:

	TUTTI I GIORNI	UNA VOLTA A SETTIMANA	MENO DI UNA VOLTA A SETTIMANA	MAI
Leggere libri insieme	☐	☐	☐	☐
Guardare la TV insieme	☐	☐	☐	☐
Guardare video insieme	☐	☐	☐	☐
Giocare insieme	☐	☐	☐	☐
Raccontare storie	☐	☐	☐	☐
Cantare canzoni o dire poesie	☐	☐	☐	☐
Ascoltare musica	☐	☐	☐	☐

5. Durante le attività insieme, con quale frequenza ...

	SEMPRE	SPESSO	A VOLTE	MAI
Parlate di quello che state facendo	☐	☐	☐	☐
Il bambino fa domande	☐	☐	☐	☐
Il bambino crea storie nuove	☐	☐	☐	☐
Il bambino ripete quello che sente	☐	☐	☐	☐
Commentate i disegni del libro	☐	☐	☐	☐
Commentate la copertina del libro	☐	☐	☐	☐

6. Con quale frequenza il bambino/la bambina fa queste attività:

	TUTTI I GIORNI	UNA VOLTA A SETTIMANA	MENO DI UNA VOLTA A SETTIMANA	MAI
Disegnare	☐	☐	☐	☐
Guardare la TV	☐	☐	☐	☐
Guardare video	☐	☐	☐	☐
Fare lavori artistici	☐	☐	☐	☐
Far finta di scrivere	☐	☐	☐	☐
Far finta di leggere	☐	☐	☐	☐
Giocare da solo	☐	☐	☐	☐
Giocare con qualcuno	☐	☐	☐	☐

7. Con quale frequenza il genitore fa queste attività:

	TUTTI I GIORNI	UNA VOLTA A SETTIMANA	MENO DI UNA VOLTA A SETTIMANA	MAI
Insegnare a contare	☐	☐	☐	☐
Insegnare a scrivere qualche parola	☐	☐	☐	☐
Insegnare a leggere qualche parola	☐	☐	☐	☐
Insegnare le lettere dell'alfabeto	☐	☐	☐	☐
Insegnare qualche parola nuova	☐	☐	☐	☐

8. Quanti libri avete a casa?

☐ Nessuno o meno di 10

☐ una mensola (10-25 libri) 

☐ uno scaffale (25-100 libri) 

☐ due scaffali o più (+ 100 libri) 

Questionnaire for teachers/educators

QUESTIONARIO INSEGNANTE

1. Compilare la tabella con i dati richiesti

	BAMBINO/A
Nome e Cognome	
Sesso	☐ **M** ☐ **F**
Data di nascita	
Luogo di nascita	
Data di arrivo a scuola	
Classe	

2. Il bambino/la bambina segue un percorso specifico di alfabetizzazione in italiano?

   ☐ SI ☐ NO

3. Il bambino/la bambina ha frequentato l'asilo nido in Italia?

   ☐ SI ☐ NO ☐ Non ha frequentato l'asilo nido ☐ Non so

4. Il bambino/la bambina ha frequentato tutta la scuola d'infanzia in Italia?

   ☐ SI ☐ NO ☐ Non so

5. Che lingue usa e sente il bambino/la bambina?

	CHE LINGUA PARLA CON …	IN CHE LINGUA PARLA AL BAMBINO…
la MADRE		
il PADRE		
FRATELLO/SORELLA		
AMICI DELLA STESSA ORIGINE		
AMICI ITALIANI		
MAESTRE		
ALTRI FAMILIARI		
ALTRI MEMBRI DELLA COMUNITÀ SCOLASTICA		

6. Che livello di italiano hanno i genitori del bambino/della bambina

	MADRE				PADRE			
	0 per niente	1 poco	2 bene	3 molto bene	0 per niente	1 poco	2 bene	3 molto bene
PRODURRE								
COMPRENDERE								

APPENDIX C

Distribution of forms in identifiable contexts by sub-type of context and by language group

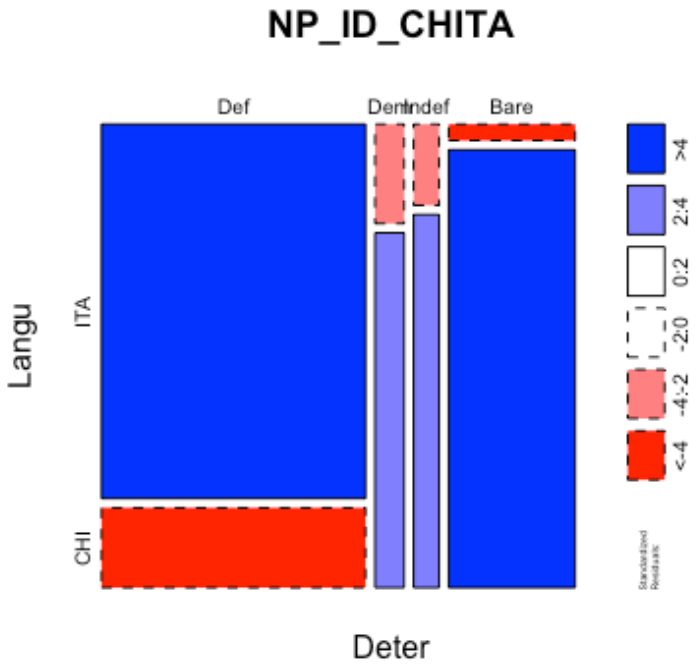
	ITA	ITA-ARA	ITA-CHI	ITA-SPA	Total
anaphoric	**192 (75%)**	**180 (74%)**	**169 (72%)**	**199 (75%)**	**740 (74%)**
definite	181 (94%)	161 (89%)	29 (17%)	190 (95%)	561 (76%)
demonstrative	7 (4%)	10 (6%)	22 (13%)	1 (0,5%)	40 (5%)
indefinite	4 (2%)	4 (2%)	23 (14%)	1 (0,5%)	32 (4%)
bare nouns	0 (0%)	5 (3%)	95 (56%)	7 (4%)	107 (14%)
associative	**37 (15%)**	**24 (10%)**	**31 (14%)**	**38 (14%)**	**130 (13%)**
definite	37 (100%)	24 (100%)	13 (42%)	34 (89%)	108 (83%)
demonstrative	0 (0%)	0 (0%)	1 (3%)	0 (0%)	1 (1%)
indefinite	0 (0%)	0 (0%)	0 (0%)	1 (3%)	1 (1%)
bare nouns	0 (0%)	0 (0%)	17 (55%)	3 (8%)	20 (15%)
generale knowledge	**26 (10%)**	**40 (16%)**	**34 (14%)**	**27 (10%)**	**127 (13%)**
definite	21 (81%)	34 (87%)	9 (26%)	24 (89%)	88 (70%)
demonstrative	0 (0%)	0 (0%)	1 (3%)	0 (0%)	1 (1%)
indefinite	0 (0%)	1 (3%)	0 (0%)	0 (0%)	1 (1%)
bare nouns	5 (19%)	4 (10%)	24 (70%)	3 (11%)	37 (29%)
situational	0 (0%)	0 (0%)	1 (0,4%)	0 (0%)	**1 (0,1%)**
demonstrative	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)
Total	**255 (100%)**	**244 (100%)**	**235 (100%)**	**264 (100%)**	**998 (100%)**

Mosaic-plots representing residual analysis

An analysis of standardized residuals was conducted to better understand the nature of differences in determiner selection between the bilingual groups and the monolingual control group. The residuals represent the discrepancies between observed values and expected values under the null hypothesis of independence between the "linguistic group" variable and determiner selection. In other words, the analysis of residuals allowed us to understand which type of determiner contributed most to the significance of the difference in determiner selection between a bilingual group and the control group. It is important to note that if a standardized residual value (SD) is greater than 1.96 or smaller than -1.96, the determiner makes a statistically significant contribution to the obtained χ2-statistic value at

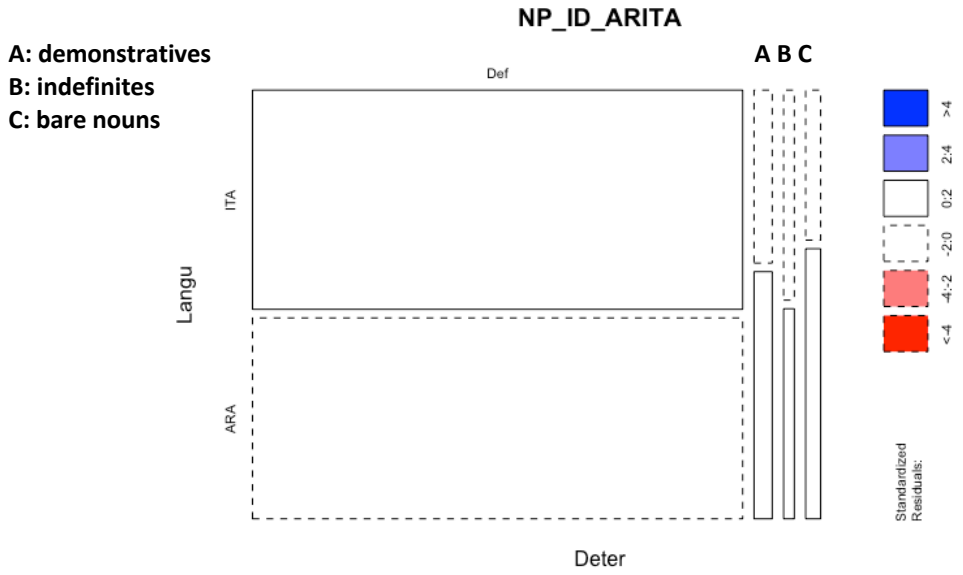
the significance level of 0.05 (Levshina, 2015). Graphical representation of residuals through a mosaic plot was chosen. Below are the mosaic plots for each comparison in the two contexts. The side legend indicates the association between the colors and the value of the standardized residuals.

1. Mosaic-plot comparing ITA and ITA-CHI's uses of determiners in identifiable contexts

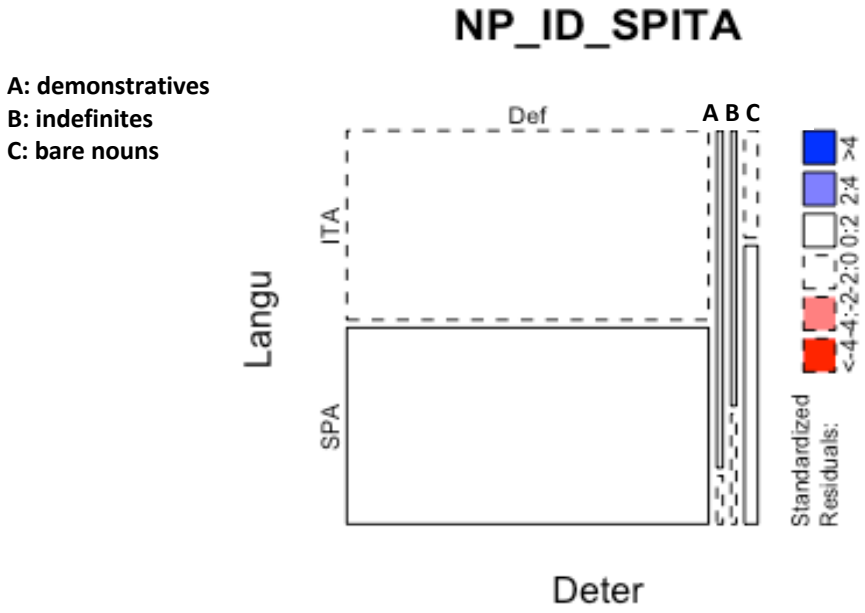


In mosaic plot 1, it is shown that in the productions of ITA-CHI bilinguals bare nouns are strongly represented (SD > 4) compared to Italian monolingual productions. This suggests that the observed differences in determiner selection are primarily attributable to the frequency of bare nouns usage in the two groups. In identifiable contexts, other determiners also contribute significantly.

2. Mosaic-plot comparing ITA and ITA-ARA's uses of determiners in Identifiable contexts

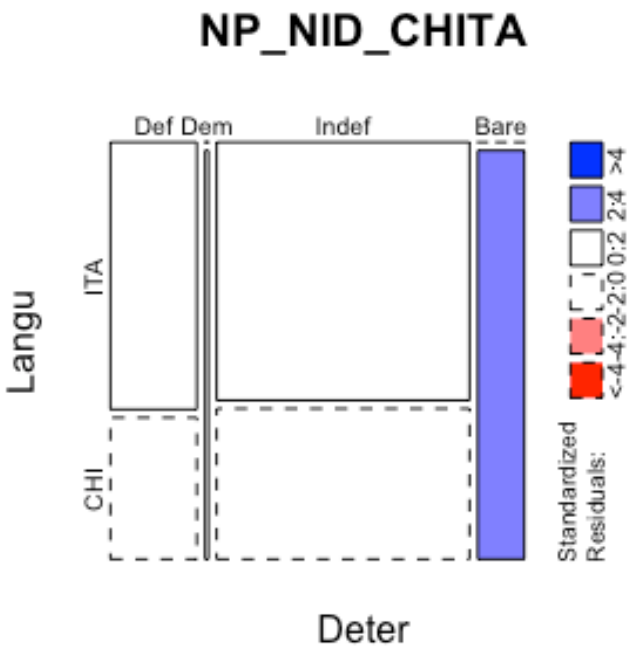


3. Mosaic-plot comparing ITA and ITA-SPA's uses of determiners in Identifiable contexts



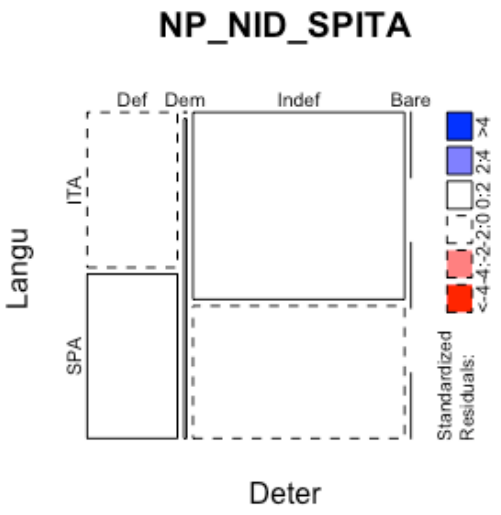
For Arabic and Spanish speakers (mosaic plots 3 and 4), no major differences in the use of determiners in identifiable contexts between monolinguals and the other two bilingual groups emerge from the analysis of residuals.

4. Mosaic-plot comparing ITA and ITA-CHI's uses of determiners in non-identifiable contexts

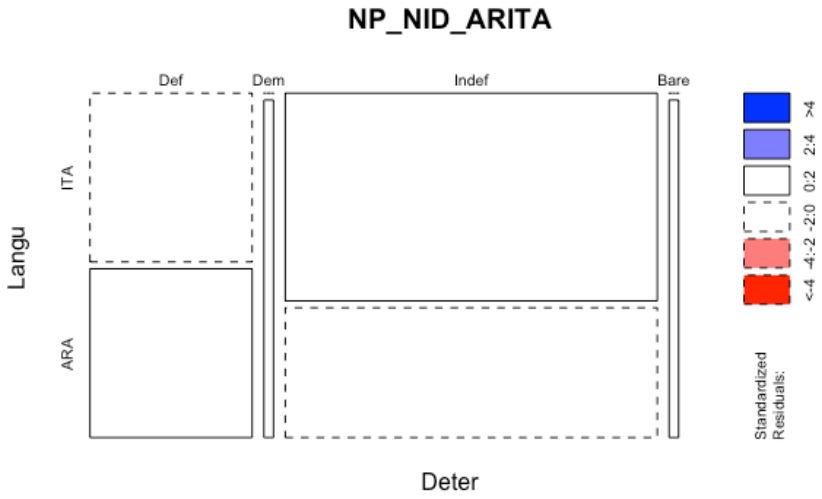


The mosaic plot 4 shows that the significance of differences between the ITA-CHI group and the monolingual control group in the selection of determiner in non-identifiable contexts is primarily due to bare nouns (2 > SD < 4), more prevalent in the bilingual group.

5. Mosaic-plot comparing ITA and ITA-SPA's uses of determiners in non-identifiable contexts



6. Mosaic-plot comparing ITA and ITA-ARA's uses of determiners in non-identifiable contexts



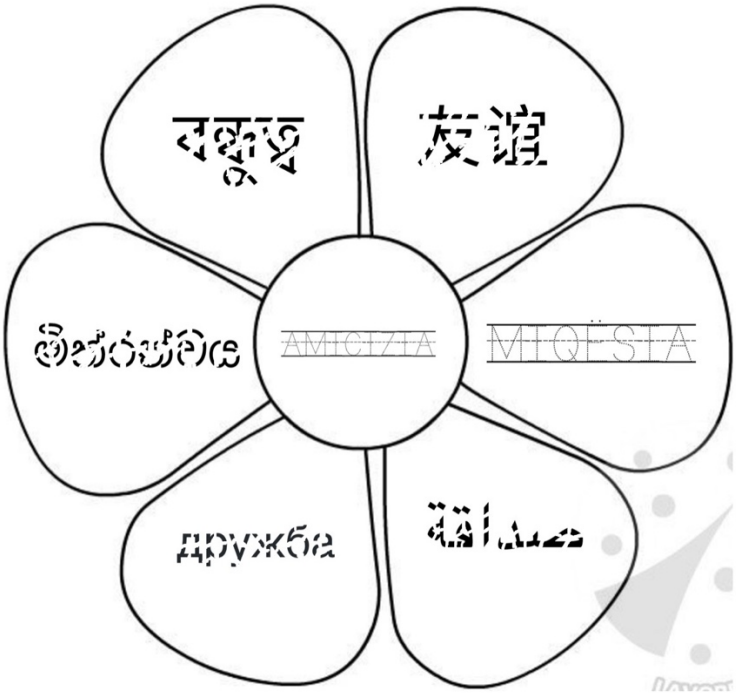
For Spanish and Arabic speakers (mosaic-plots 5 and 6), no major differences in the use of determiners in non-identifiable contexts between monolinguals and the other two bilingual groups emerge from the analysis of residuals.

Linguistic profile in Family Language, by language

	ITA-CHI			ITA-SPA			ITA-ARA		
(max)	**Range**	**Mean**	**sd**	**Range**	**Mean**	**sd**	**Range**	**Mean**	**sd**
MLU5 (20.80)	17.40-10.60	13.50	2.8	20.80-9.60	16.12	3.5	10.20-2.60	7.14	2.7
Sub-Index (1.50)	1.29-0.78	0.96	0.1	1.50-0.51	1.17	0.3	1.33-0.48	0.88	0.3
VOCD (49.93)	49.93-21.71	32.12	10	45.28-14.02	27.06	9.1	41.08-16.21	24.83	8.6
NDV (35)	35-19	26.90	6.9	29-7	16.10	7.9	24-2	13.40	7.7

APPENDIX D

- **Task 3, Step 2, Corner 3: flower with dotted scripts**



- **Task 4: text of the story "La pecorella vanitosa" in Italian (with English version), provided as starting point for translations**

LA PECORELLA VANITOSA (ITALIAN)

[1] Questa è la storia di una pecorella vanitosa, che adorava farsi guardare da tutta la fattoria. La pecorella era stanca del suo pelo bianco. "Mhmm che noia, così tutta bianca nessuno capirà quanto sono bella!" diceva ogni giorno. Ed era convinta che con la lana di un altro colore tutti l'avrebbero ammirata ancora di più. Decise allora di colorarsi tutta.
[2] Si buttò in mezzo all'erba e iniziò a rotolarsi e fare capriole e come per magia diventò tutta verde!
[3] In quel momento arrivò un toro di nome Gedeone, grande, grosso e un po' tontolone, che vedendo la pecorella, la scambiò per un cespuglio d'erba. "Gnam gnam" disse il toro e si chinò per mangiarla. Via! … la povera pecora scappò di corsa e pensò che il verde non fosse il colore adatto a lei.
[4] "Sarò più bella con il pelo rosso!" disse e proprio in quel momento vide un cesto pieno di fragole. Decise allora di buttarcisi dentro. Fece un bel bagno in quel cesto e ne uscì tutta coperta di polpa color rosso vivo … sembrava proprio un fragolone gigante!
[5] Quando Lina, la contadina, vide la pecorella, la scambiò per una fragola gigante e la rincorse per farci una buonissima marmellata! Via! … la povera pecora scappò di corsa e pensò che il rosso non fosse il colore adatto a lei.

THE LITTLE VAIN SHEEP (ENGLISH)

[1] This is the story of a little vain sheep who loved to be looked at by the whole farm. The little sheep was tired of her white wool. "Hmm, so boring, all white as I am, no one will understand how beautiful I am!" she said every day. She was convinced that with wool of another color, everyone would admire her even more. So, she decided to color herself all over.
[2] She jumped into the grass and started rolling and doing somersaults, and magically, she turned all green!
[3] At that moment, a bull named Gideon came along. He was big, bulky, and a bit clumsy. When he saw the little sheep, mistook her for a clump of grass. "Yum yum!" said the bull, and he bent down to eat her. Zoom! …The poor sheep ran off in a hurry, thinking that green wasn't the right color for her.
[4] "I'll be prettier with red wool!" she said, and just then, she saw a basket full of strawberries. So, she decided to jump in. She took a nice bath in that basket and came out all covered in bright red pulp... she really looked like a giant strawberry!
[5] When Lina, the farmer, saw the little sheep, she mistook her for a giant strawberry and chased her to make delicious jam! Zoom!... the poor sheep ran off in a hurry, thinking that red wasn't the right color for her.

[6] “Sarò più bella con il pelo giallo!” pensò e proprio lì vicino vide un alveare con del miele giallo come l’oro, che brillava alla luce del sole. “Sarò la più bella della fattoria con il pelo dorato!” disse e così corse a spalmarsi addosso quel miele, ma … appena le api la videro iniziarono a inseguirla! Via! … la povera pecora scappò di corsa e pensò che il giallo non fosse il colore adatto a lei.
[7] “Sarò più bella con il pelo marrone!” pensò e si buttò in un vaso di cioccolata.
[8] Un bambino e una bambina che vivevano nella fattoria la videro e la scambiarono per Nutella. “Mhmm che merenda deliziosa faremo oggi con tutta questa Nutella!” dissero e iniziarono a inseguirla. Via! … la povera pecora scappò di corsa e pensò che il marrone non fosse il colore adatto a lei.
[9] “Ora basta, non ne posso più, rivoglio il mio colore!” urlò forte la pecorella, stanca di essere inseguita. E così se ne andò a riposare sotto gli alberi e capì che non c’era bisogno di cambiare colore per essere bella, perché tutti siamo belli proprio così come siamo.

[6] “I'll be prettier with yellow wool!” she thought, and right nearby, she saw a beehive with yellow honey as golden as the sun. “I’ll be the most beautiful of the farm with golden wool!” she said and ran to smear that honey all over herself, but... as soon as the bees saw her, they started chasing her! Zoom!... the poor sheep ran fast and thought that yellow wasn't the right color for her.
[7] “I’ll be prettier with brown wool!” she thought and jumped into a pot of chocolate.
[8] A boy and a girl who lived on the farm saw her and mistook her for Nutella. “Yum, what a delicious snack we’ll have today with all this Nutella!” they said and started chasing her. Zoom!... the poor sheep ran fast and thought that brown wasn't the right color for her.
[9] “Enough now, I can’t take it anymore, I want my color back!” the little sheep shouted, tired of being chased. And so, she went to rest under the trees and understood that there was no need to change color to be beautiful because we are all beautiful just as we are.

- **Task 4: combination of languages and story**

[1]: Italian; [2]: Bengali/English; [3]: Romanian; [4] Spanish; [5] Mandarin Chinese; [6] Moroccan Arabic/Standard Arabic; [7] Filipino; [8] Urdu/English; [9] Albanian/Tigrinya

- **Task 4: transcriptions of the audio in different family languages from “La pecorella vanitosa”**

[2] **Bengali:**

ভেড়াটি ঘাসের মধ্যে নিয়ে পড়ে গেল গড়াগড়ি করল জাদুতে সবুজ হয়ে গেল।

[3] **Romanian**:

În acel moment a susțit un taur pe nume Gideon, mare, mare și puțin prostuț, care, văzând oaia, a confundat-o cu un tufiș. “Iam, iam!” a spus taurul și s-a plecat să înceapă să o mănânce. Beata oaie a început să fugă, realizând că verdele nu este culoarea potrivită pentru ea.

[4] **Spanish**:

“¡Estaré más guapa con el pelo rojo!” dijo la ovejita, y justo en ese momento vio una cesta llena de fresas. Al verla, decidió tirarse dentro y se bañó inmersa en las fresas. Después de un rato, salió toda cubierta de pulpa de fresa. Estaba totalmente roja... ¡parecía un fresón gigante!

[5] **Mandarin Chinese:**

麗娜農人看到了小羊之後,因為覺得小羊跟一個很大的草莓一模一樣,就開始追著它,想把它作為一種非常好吃的果子醬。

[6] **Moroccan Arabic**

وذهبت .الدهبي بالشعر كلها المزرعة في أجمل نكون سوف الدهب بحل الصفر العسل في نحل عشرة رأيت منها قريب ومن ما باللصفر تفكر وبدأت .بسرعة تجربة بدأت المسكينة والمعيذة .تباعها بدأت نحل رأيت عندما ولكن :.فوقها العسل لتدهن لتجربة .لها المناسب اللون هي

[6] **Standard Arabic:**

.الشمس ضوء تحت يلمع كان ذهبي أزفر لون عسل في نحل خلية رأت منها قريب مكان وفي !الأزفر بالشعر أجمل سأكون مطاردتها في بدأوا النحل رآها عندما ولكن ،جسمها على العسل تضاع لأن فركضت !الذهبي بالشعر المزرعة في الأجمل سأكون .لها المناسب اللون ليس الأزفر اللون أن واعتقدت ،المسكينة النعجة هربت

[7] **Filipino:**

Mas gaganda ako sa brown na buhok. Naisip niya at itinapon ang sarili sa isang kaldero ng tsokolate.

[8] **Urdu:**

فارم پر کھڑے لڑکا اور لڑکی نے جب بھیڑ کو دیکھا تو وہ اس کو نیوٹیلا سمجھ بیٹھے انہوں نے کہا آج کتنے مزے کا کھانا کھائیں گے نیوٹیلا کے ساتھ اور اس کا پیچھا کرنے لگے بیچاری بھیڑ بھاگ پڑی اور اس نے سوچا یہ بھورا رنگ میرے لئے نہیں ہے

[9] **Albanian:**

Tani boll nuk duroi dot më dua të më rikthehet ngjyra delja bërtiti me zë të lartë e lodhur nga ndjekja. Ajo shkoi të pushonte nën pemë. Dhe kuptoi se nuk kishte nevojë të ndryshonte ngjyrën për të qenë e bukur, sepse të gjithë jemi të bukur ashtu siç jemi.

[9] **Tigrinya**:

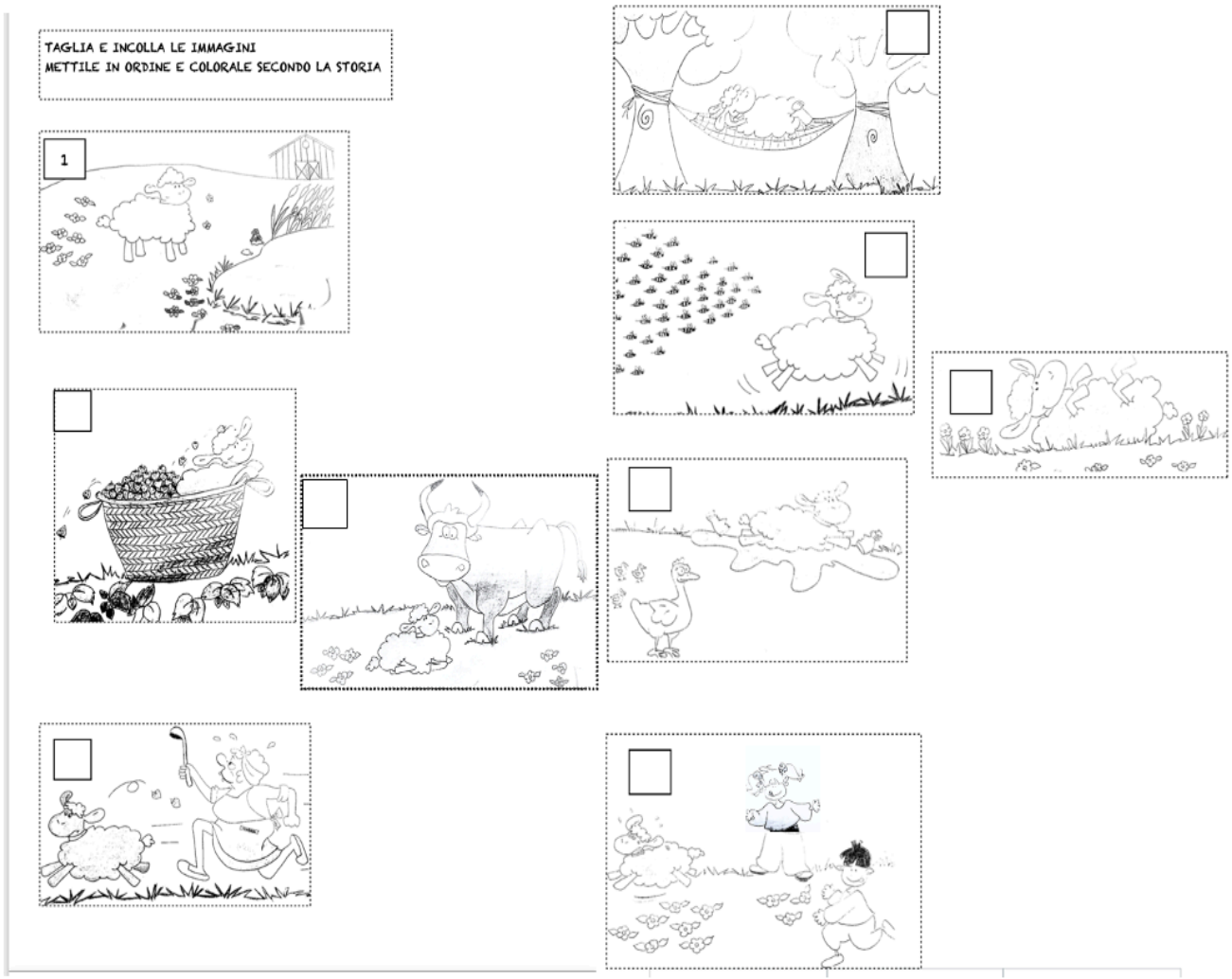
እታ ንእሽተይ ማሕስእ (በጊዕ) ነያ ነያ ምስ ደኸመት“ይኣክል ድሕሪ ሕጂ ይኣኽለኒ በቃ፡ናብሕብረይ ክምለስ እየ ዝደሊ " ኢላ ጨደረት። ኣብ ትሕቲ ኣግራብ ከተዕርፍ ከደት፡ ጽብቕቲ ክትከውን ሕብሪ ምቕያር ክምዘየድሊ ተረድኣት፡ ምኽንያቱ ኩላትና ልክዕ ክምቲ ዘለናዮ ጽቡቓት ኢና።"

All transcriptions were created based on the audio sent by the native speakers contacted for Task 4.

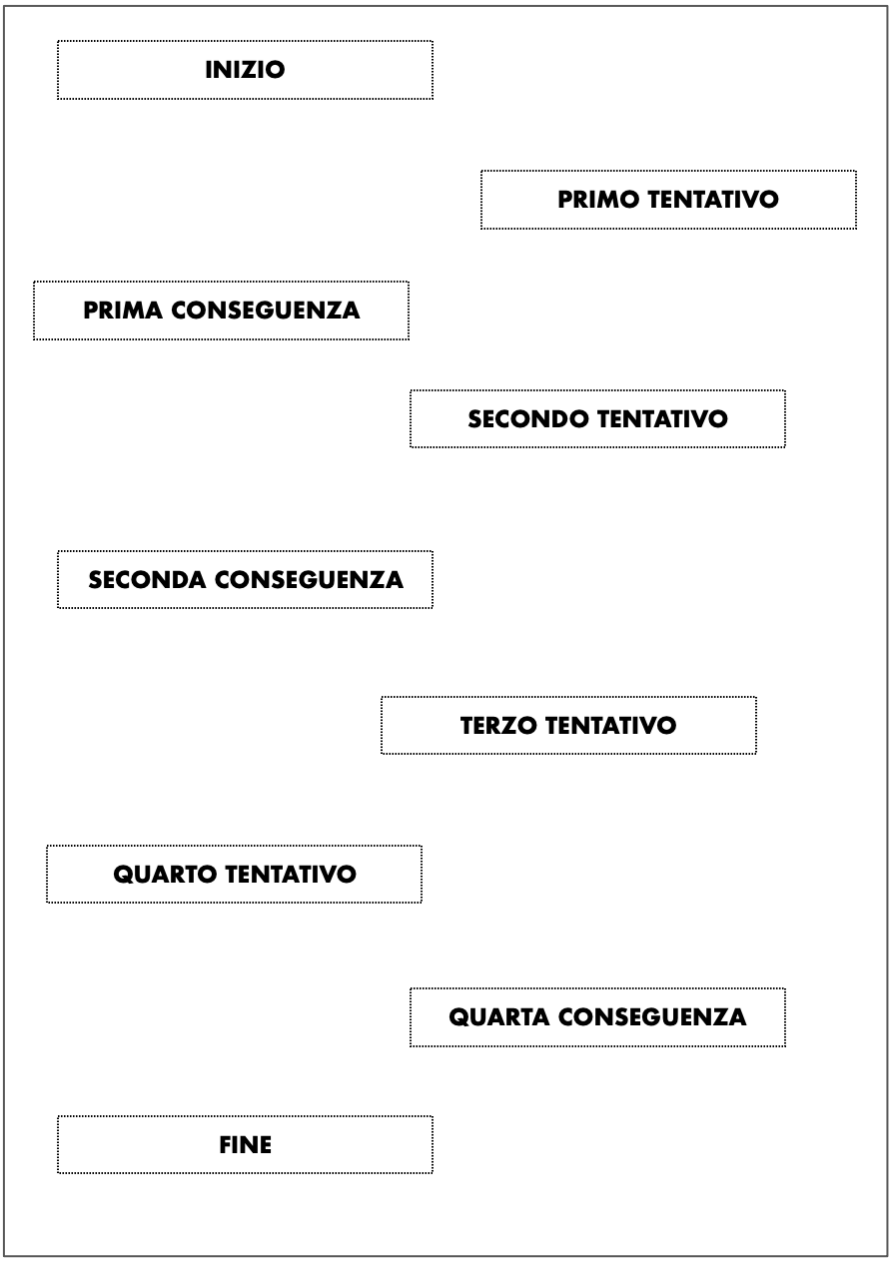
Two freely available online automatic speech-to-text tools were used:

https://voiser.net; https://app.soundtype.ai

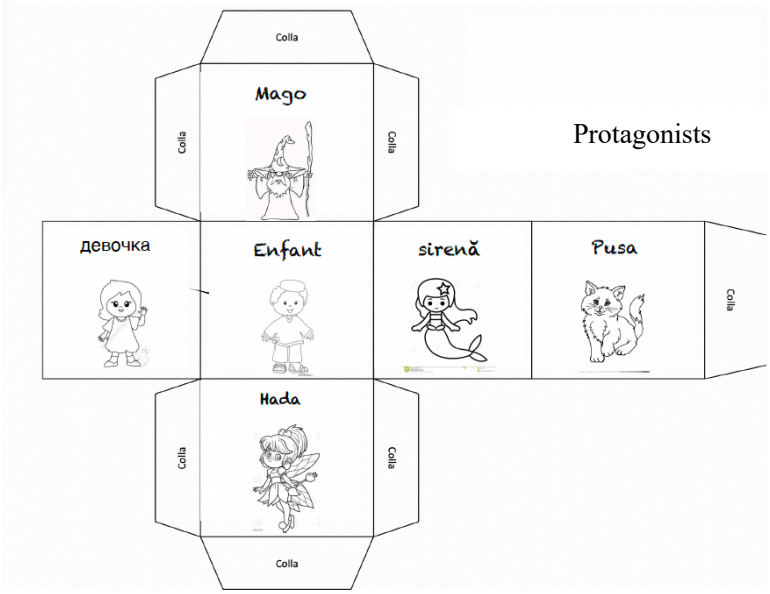
- **Task 4: worksheet on the story “La pecorella vanitosa” (sequences to be ordered) and coloured**



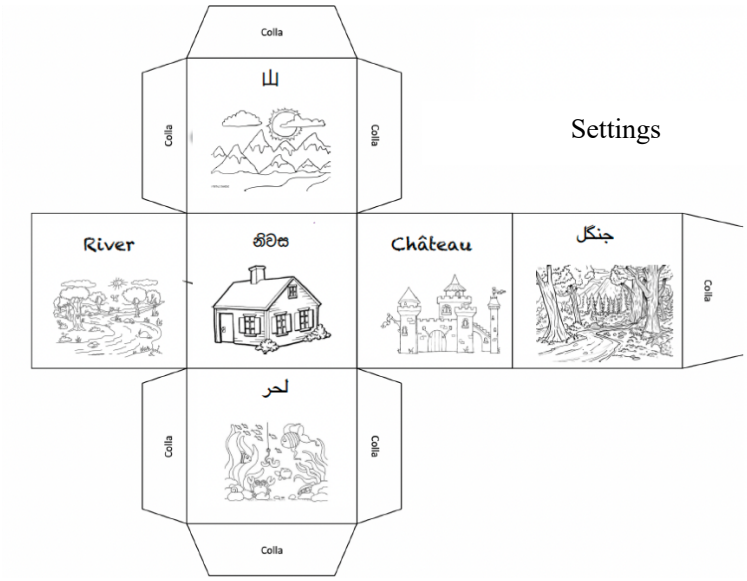
- **Task 4: worksheet on the story "La pecorella vanitosa" (grammar story labels)**



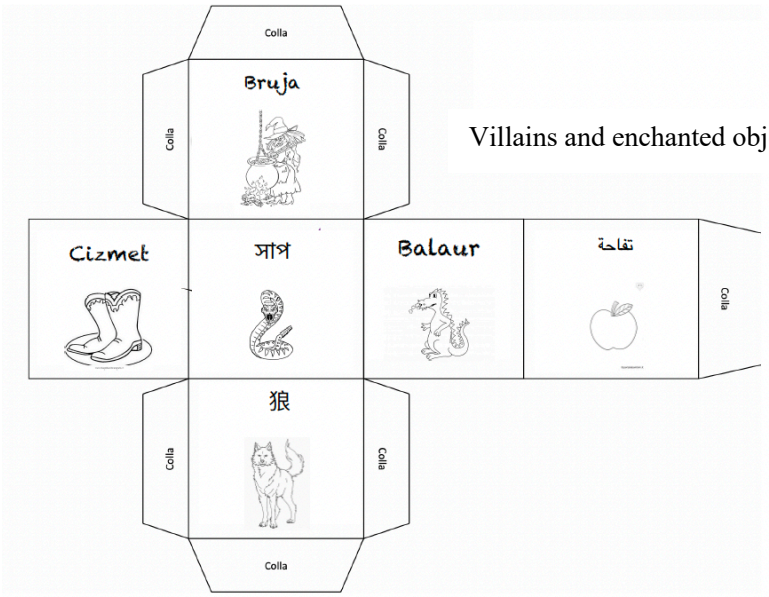
- **Task 5: storytelling dice**



Protagonists



Settings



Villains and enchanted objects

- **Task 5: opening formulas**

C'ERA UNA VOLTA

Na ishte një herë...

.....

كان يا ما كان،في قديم الزمان، وسالف العصر والأوان

kan ya ma kan, fi qadim alzaman, wasalif aleasr wal'awan

.....

এককালে

Ēkakālē

.....

很久很久以前

Hěnjiǔ hěnjiǔ yǐqián

.....

Noong unang panahon

.....

Il était une fois

.....

Once upon a time

.....

- **Task 5: Example of a final story**

C'ERA UNA VOLTA ...

UN MAGO CHE AVEVA POTERI STRAORDINARI: POTEVA TELETRASPORTARSI DALL'ALTRA PARTE DEL MONDO IN TRE SECONDI E DIVENTARE INVISIBILE. AVEVA UN CAPPELLO APPUNTITO, UN VESTITO MOLTO LUNGO E UN BASTONE CON CUI FACEVA TUTTE LE SUE MAGIE.

UN **CASTILLO** *(spa.castello)* ERA LA SUA **HOUSE** *(ing.casa)*.

UN GIORNO, UN **শিশু ŚIŚU** *(bang. bambino)* ANDÒ DAL MAGO E GLI CHIESE UNA POZIONE MAGICA PER REALIZZARE IL SUO **梦想 MÈNG XIĂNG** *(cin. sogno)*, CHE ERA QUELLO DI VIVERE IN UN GRANDE **CASTILLO**, VISTO CHE LA SUA **বাড়ি BĀṚI** *(bang. casa)* ERA **MUY PEQUEÑA** *(spa.molto piccola)*.

IL MAGO, PERÒ, PREPARÒ LA POZIONE SBAGLIATA, PERCHÉ AVEVA SCAMBIATO GLI INGREDIENTI CORRETTI CON ALTRI CHE NON C'ENTRAVANO NIENTE: **梨 LÍ** *(cin. pera)* AL POSTO DEL KIWI E **PIÑA** *(spa.ananas)* AL POSTO DELLA BANANA.

QUANDO IL MAGO SI ACCORSE DI AVER SBAGLIATO, **EL CHICO** *(spa.bambino)* SE NE ERA GIÀ ANDATO, TUTTO CONTENTO DI AVER REALIZZATO IL SUO **梦想 MÈNG XIĂNG** ED ERA ORMAI TROPPO LONTANO. IL MAGO, ALLORA, DECISE DI RAGGIUNGERLO CON IL TELETRASPORTO, MA QUANDO ARRIVÒ VIDE CHE IL **শিশু ŚIŚU** AVEVA GIÀ PROVATO LA POZIONE E LA **বাড়ি BĀṚI** ERA DIVENTATA ANCORA PIÙ PICCOLA! IL **ДИТИНИ DYTYNY** *(ucr.bambino)* MOLTO **ARRAGGIAT** *(nap. arrabbiato)*, ORDINÒ AL MAGO DI PREPARARE SUBITO UNA NUOVA POZIONE.

TORNATO AL CASTELLO, IL MAGO CI RIPROVÒ, MA SBAGLIÒ ANCORA UNA VOLTA E LA دار **DAAR** *(ara.casa)* DEL BAMBINO DIVENTÒ PIÙ **PEQUEÑA** DI UNA FORMICA!
IL MAGO CONTINUAVA A SBAGLIARE, ERA TROPPO DISTRATTO! PENSÒ ALLORA DI CONTROLLARE LA RICETTA DELLA POZIONE PER CAPIRE COSA SI STESSE DIMENTICANDO, E SOLO IN QUEL MOMENTO SI RICORDÒ DI AVERLA LASCIATA NELLA GROTTA DI UN **URS** *(rum.orso)*. DECISE ALLORA DI TELETRASPORTARSI NELLA GROTTA, MA QUANDO ARRIVÒ TROVÒ LA RICETTA STRAPPATA IN TANTI PICCOLI PEZZI! OH NO, CHE DISDETTA! MA PER FORTUNA IL MAGO AVEVA I SUOI POTERI E FECE APPARIRE LO SCOTCH. COSÌ, CON PAZIENZA, MISE INSIEME TUTTI I PEZZETTI DI CARTA E RICOSTRUÌ LA RICETTA.

FINALMENTE, IL MAGO RIUSCÌ A PREPARARE LA POZIONE GIUSTA PER TRASFORMARE LA دار **DAAR** DEL طـــفل **TIFL** *(ara.bambino)* IN UN GRANDE **CASTILLO**. QUESTA MAGIA, PERÒ, ERA COSÌ POTENTE CHE IL MAGO ESAURÌ TUTTE LE SUE FORZE E RIMASE SENZA POTERI, DIVENTANDO UNA PERSONA COME TUTTE LE ALTRE. AVEVA DATO TUTTO QUELLO CHE POSSEDEVA PER REALIZZARE IL 梦想 **MÈNG XIĂNG** DI UN শিশু **ŚIŚU.** IL **ДИТИНИ DYTYNY** ERA COMMOSSO E PER RINGRAZIARLO LO INVITÒ NEL SUO NUOVO CASTELLO E, INSIEME AGLI ABITANTI DEL VILLAGGIO, FESTEGGIARONO TUTTA LA NOTTE, FELICI E CONTENTI!

The story was invented by children of the primary school (second grade) of I.C. 11