Universitat Autònoma de Barcelona

STUDIES IN EARLY MANDARIN SYNTAX

JINGTAO ZHU

STUDIES IN EARLY MANDARIN SYNTAX

JINGTAO ZHU

Supervised by Dr. Anna Gavarró

Universitat Autònoma de Barcelona

A doctoral dissertation submitted for the degree of
Philosophiæ Doctor in Cognitive Science and Language

Facultat de Filosofia i Lletres
Departament de Filologia Catalana
Centre de Lingüística Teòrica

September 2021

ABSTRACT

While many studies address the emergence of phonological awareness from birth, little is known about the development of syntax in infancy. In this dissertation I offer a corpus-based study and two experiments on the development of syntax in Mandarin Chinese and present novel results that challenge lexical approaches to language acquisition (e.g., Abbot-Smith & Tomasello, 2006; Tomasello, 2003) and also the role of frequency. The corpus study based on the spontaneous speech of 47 typically-developing, monolingual-Mandarin speaking children (between the age of 1;2 to 6;5) shows that their performance on both null subjects and null objects has already become adult-like by 1;8, and children produce the non-canonical S*ba*OV construction at least when they approach their second birthday (at 1;8) and demonstrate productive knowledge at age 3-4. The relatively high frequency of null objects in Mandarin and the low frequency of the *ba* construction in the input should lead children to drop objects (i.e., post-*ba* NPs) in this construction at least at the early stage if frequency determines the path of acquisition. However, this prediction is not fulfilled: even at age 1;8 children produce post-*ba* NPs in the *ba* construction consistently, and this has implications against variational accounts of language acquisition (Yang, 2002, 2004). The first experiment replicates that of Franck et al. (2013) for French; with eye-tracking techniques, I compare well-formed transitive SVO sentences with ill-formed SOV sequences, all involving pseudo-verbs, when presented to infants with a mean age of 17.4 months. The results show that infants comprehended only when exposed to a grammatical sentence. Neither age nor vocabulary are predictors of performance. In the second experiment replicating Lassotta et al. (2014), ran with 24 infants with mean age 17.5 months, infants were confronted with canonical SVO and non-canonical S*ba*OV and OS*ba*OV sentences. The analysis indicates that hearing the SVO sentences causes infants to look significantly longer at the scene with the first NP as AGENT, and the same happens when they hear the S*ba*OV sentences. However, infants exhibit an opposite eye movement pattern when they hear the OS*ba*OV sentences, with more fixations on the scene with the first NP as PATIENT, indicating comprehension of all these grammatical sequences. Combining the results of the three studies I conclude that Mandarin infants are sensitive to word order at a very early age, 17.5 months at the latest. In fact, they are able to parse canonical and non-canonical word orders (the *ba* construction with a preverbal object, appearing either before or after the subject) which are scarce in the input, while they ignore ill-formed sequences. Therefore, these results can only be accounted for if grammatical, language-specific knowledge is available, and constitute evidence for very early parameter setting.

RESUM

Tot i que hi ha molts estudis sobre l'aparició de la sensibilitat a la fonologia des del naixement, no n'hi ha tants sobre el desenvolupament de la sintaxi en la primera infantesa. En aquesta tesi faig un estudi de corpus i dos experiments sobre el desenvolupament de la sintaxi en xinès mandarí i presento resultats nous que qüestionen les aproximacions lèxiques a l'adquisició (Abbot-Smith & Tomasello, 2006; Tomasello, 2003) i les aproximacions basades en la freqüència. L'estudi de corpus es basa en l'anàlisi de les produccions espontànies de 47 nens monolingües amb desenvolupament típic d'edats compreses entre 1;2 i 6;5 anys; s'analitza la producció de subjectes i objectes i s'observa que la producció de subjectes i objectes nuls és adulta a l'edat de 1;8; els nens produeixen la construcció S*ba*OV a l'edat d'1;8 i demostrem que en fan un ús adult a l'edat de 3-4 anys. La freqüència relativament alta d'objectes nuls en mandarí i la freqüència baixa de les frases amb ba podria fer esperar l'omissió dels objectes a la posició posterior a ba, almenys en un primer estadi; però aquesta predicció no es compleix: fins i tot a l'edat d'1;8 els infants produeixen l'objecte posterior a *ba* com a fonèticament ple, cosa que té implicacions per a la proposta variacional de Yang (2002, 2004). El primer experiment replica l'experiment de Franck et al. (2013) del francès; amb tècniques de monitorització de la mirada comparo la comprensió de la frase ben-formada SVO amb la de la frase mal-formada SOV, amb l'ús de pseudo-verbs. Els infants, amb una edat mitjana de 17,4 mesos, van comprendre la frase només quan era ben-formada. Ni l'edat ni la mida del vocabulari no podien predir els nivells de comprensió. Al segon experiment que replica l'experiment de Lassotta et al. (2014), els nens sentien frases amb l'ordre canònic SVO i frases amb ordres no-canònics S*ba*OV i OS*ba*OV. L'anàlisi dels resultats indica que els infants dirigien la mirada a l'escena amb un AGENT en primera posició quan sentien les frases SVO i S*ba*OV, però no quan sentien una frase OS*ba*OV, cosa que indica que eren capaços d'entendre aquests tres tipus de frases gramaticals. Combinant els resultats dels tres estudis, arribo a la conclusió que els infants exposats al xinès mandarí han establert l'ordre de paraules molt aviat, com a molt tard al 17,5 mesos. Saben analitzar frases amb ordre canònic i no canònic (en concret la construcció amb *ba* amb l'objecte abans o després del subjecte) molt infreqüents, mentre ignoren les seqüències mal-formades. Aquests resultats només són explicables si els infants tenen accés a coneixements gramaticals i específics de la llengua que adquireixen i que advoquen a favor de la fixació molt primerenca dels paràmetres.

RESUMEN

A pesar de que existe una gran cantidad de estudios que abordan la aparición de la sensibilidad fonológica desde el nacimiento, no hay muchos que aborden el desarrollo de la sintaxis durante la primera infancia. En esta tesis presento un estudio de corpus y dos experimentos acerca del desarrollo del orden de la sintaxis en el chino mandarín; además presento resultados novedosos que cuestionan los enfoques léxicos del desarrollo del lenguaje (Abbot-Smith & Tomasello, 2006; Tomasello, 2003) y el papel que juega la frecuencia. El estudio de corpus se basa en las producciones espontáneas de 47 niños monolingües con un desarrollo normal de edades comprendidas entre 1;2 y 6;5 años; se analiza la producción de sujetos y objetos y se observa que la producción de sujetos y objetos nulos es adulta a la edad de 1;8. Los niños producen la construcción S*ba*OV a los 1;8 y evidencian un conocimiento productivo a la edad de 3-4 años. La frecuencia relativamente alta de objetos nulos en mandarín y la frecuencia baja de las frases con *ba* debería llevar a los niños a producir nulos objetos en posición posterior a *ba*, al menos en la etapa inicial en caso de que la frecuencia determine la trayectoria del aprendizaje; no obstante, esta predicción no se cumple dado que, incluso en la edad 1;8, los niños producen los objetos posteriores a *ba* en forma fonéticamente plena, lo que tiene implicaciones para la propuesta variacional de Yang (2002, 2004). El primer experimento replica el de Franck et al. (2013) del francés; con técnicas de monitorización de la mirada, comparo oraciones transitivas SVO –formadas correctamente– con secuencias SOV –formadas incorrectamente-, todas ellas con pseudo-verbos. Los niños, con una edad media de 17,4 meses, comprendieron las oraciones solo cuando su estructura era grammatical. Ni la edad ni el vocabulario eran predictores del comportamiento. En el segundo experimento que replica el de Lassotta et al. (2014), realizado con niños con una edad media de 17,5 meses, se presentaron oraciones con el orden canónico SVO y oraciones con los órdenes no canónicos S*ba*OV y OS*ba*OV. El análisis de los resultados indica que los niños dirigen la mirada a la escena con un AGENTE en primera posición cuando escuchan las oraciones SVO y S*ba*OV, pero no cuando escuchan una oración OS*ba*OV, lo cual es un indicio de compresión de todas estas secuencias gramaticales. Combinando los resultados de los tres estudios, se concluye que los niños que tienen el mandarín como lengua materna son sensibles al orden de las palabras desde una edad muy temprana, como muy tarde a los 17,5 meses. Saben analizar frases con orden canónico y con órdenes no canónicos infrecuentes (concretamente la construcción *ba* con un objeto preverbal, el cual aparece antes o después del sujeto), mientras que ignoran las secuencias agramaticales. Estos resultados solo pueden explicarse si disponen de conocimientos gramaticales específicos de la lengua, y constituyen una prueba de que la fijación de parámetros es muy temprana.

DEDICATION

—— 纪念肥肥妈

ACKNOWLEDGEMENTS

Reading countless pages of acknowledgements in countless books, I never realized how difficult a task it would be to say "thank you" for all the nice meetings, support, help, understanding, comments and arguments given by so many people from a large and disparate community over the time it took to write this dissertation.

First and foremost, I would like to say a very big thank you to my supervisor, Prof. Anna Gavarró, who has been supportive since the days I began working on the M.A. thesis 7 years ago. Anna deserves more than words of thanks for her indispensable comments, for her patience and for all the opportunities she offered me to develop my research skills. During my entire stay at UAB, she has taught me a lot that has now become part of me. The extent to which she has made herself available to me and the amount of thought she has put into my work go well beyond what I or anyone else can reasonable expect. Her confidence in me and interest in my work has also kept me up over the years. I appreciate all her contribution of time, ideas, and funding to make my Ph.D experience productive and stimulating. I am also thankful for the excellent example she has provided as a successful researcher, professor and a wonderful woman.

During the past years, I had the good fortune to be offered the opportunity to present versions of this dissertation to audiences that never failed to provide valuable comments. It is impossible to mention everybody who had an impact to this work, but I owe a special debt to Prof. Xiaolu Yang, to Prof. Doris Chun-yin Chen, to Prof. Boping Yuan and to Prof. Carla Soares-Jesel, for their interest in the present work, their advice and constant encouragement. I am also deeply indebted to Prof. Elena Lieven, for her constructive criticism and to Prof. Thomas Hun-Tak Lee for his valuable input.

Some faculty members of other universities have been kind enough to extend their help at various phases of this dissertation as well. I cannot fail to mention Prof. Julie Franck, who discussed several aspects of the work, made valuable suggestions, provided extremely useful comments. Prof. Luigi Rizzi, who asked important questions that helped clarify my thought. I also want to thank Prof. Roberta Golinkoff, for her relevant comments on the age issue, and thank Prof. Aoju Chen and Prof. Maria del Mar Vanrell for discussion and help with the intonation of Mandarin. I am also grateful to Anna Espinal (in memoriam) and Ester Boixadera who have extended their support in statistical data analysis and to Dr. Claudia Manetti for her thoughts on the acquisition of the *ba* construction. I also thank Prof. Pilar Barbosa for drawing attention to a recent analysis of null objects.

The dissertation has also benefited from comments and valuable discussions of diverse kinds made by Prof. Stephanie Durrleman, Prof. Ken Wexler, Prof. Tom Roeper, Prof. Joel Wallenberg, Prof. Elin Thordardottir, Prof. Elena Pagliarini, Dr. Keith Tse, Yangyu Sun and anonymous reviewers from different publications. I thank them for all their contributions.

I am forever thankful for my colleagues at Centre de Lingüística Teòrica, they all have contributed immensely to my personal and professional time at UAB: Dr. Io Salmons, Dr. Celina Agostinho, Dr. Shenai Hu, Alejandra Keidel, Iman el Hadef, Tala Nazzal, Dària Serés. I want to give particular thanks to Dr. Maya Leela, who introduced me to eye-tracking; thanks for all her help in overcoming technical issues. I am also very grateful to all the teaching and research staff that have taken some time to discuss and enrich my work, especially to Prof. Jaume Mateu, Prof. Sergio Balari, Prof. Xavier Villalba, Prof. Ángel Gallego, Prof. Anna Bartra, Prof. Jaume Solà, Prof. Josep M. Brucart. I am also grateful to the administrative team, especially to Elías Gallardo and Melina Rodríguez, who kept me organized and were always ready to help.

I cannot forget my friends, who went through hard times together, cheered me on, and celebrated each accomplishment. I want to thank Aleix González, Ivan Penas, Lu Gao, Iker Mons, Mingshan Xu for giving me much help and encouragement in one way or another. I want to thank my parents for bringing me up, for all their love and for all the things that have made me always proud of them.

I am so grateful to ClicAsia, Centre d'Estudis Orientals and the project FFI2017-87699-P for financial support, without their help, I couldn't have finished so smoothly. Many thanks are due to the infants and adults who agreed to participate in the experiments reported, to nurses and nurse aids in the Danan Community Health & Services Centre, although most of these people must remain nameless in this document due to concern for their confidentiality, I must admit that without their cooperation and support, this PhD study would not have been possible.

I send my deep thanks for your contribution to this dissertation, but the inevitable mistakes in it are very much my own.

PUBLICATIONS

Some of the results in this dissertation have appeared or are to appear in the following publications:

- Zhu, J.T. & Gavarró, A. (2019). Testing language acquisition models: Null and overt topics in Mandarin. *Journal of Child Language*, 46(4), 707-732. DOI: https://doi.org/10.1017/S0305000919000114
- Gavarró, A. & Zhu, J.T. (2019/2020). Functional categories in very early acquisition. In A.M. Di Sciullo (Ed.), *Approaches to Functional Categories, Linguistic Analysis Thematic Issue*, 42(3-4), 583-599.
- Zhu, J.T. Franck, J., Rizzi, L., & Gavarró, A. (2021). Do infants have abstract grammatical knowledge of word order at 17 months? Evidence from Mandarin Chinese. *Journal of Child Language*, first view, 1-20. DOI: https://doi.org/10.1017/S0305000920000756

ETHICAL CONSIDERATIONS

Both the comprehension of canonical word order experiment (Chapter 4) and the comprehension of non-canonical word order experiment (Chapter 5) have been approved by the Comissió d'Ètica en l'Experimentació Animal i Humana of the Universitat Autònoma de Barcelona with the project title "An experiment of comprehension of word order using eye-tracking techniques in infants at 17 months" presented by Jingtao Zhu (approval number 5071). Informed consent was obtained from each family prior to the experiment by their caregivers in the case of infant participants and the informed consent from adult participants was obtained before the beginning of the experiment as well.

Contents

Abbreviations

ACC	Accusative case
ADV	Adverbial marker
Aux	Auxiliar
BA	Ba construction
BEI	Bei construction (passive)
BI	Comparative marker
C	Complementizer
CL	Clitic
CLASS	Classifier
CLE	Cleft configuration (shi...de)
Comp	Complement
DC	Directional complement
DE	Possessive modifier
DET	Determiner
DOM	Differential object marking
ERG	Ergative case
EXL	Exclamation
EXP	Experiential aspect
FOC	Focus marker
FV	Final vowel
NEG	Negation
NOM	Nominative case
ORT	Cohort modal
PAST	Simple past tense
PERF	Perfective aspect
PN	Recent past
PSEUDO-V	Pseudo-verb
Q	Polar question particle
SFP	Sentence-final particle
SM	Subject marker
Spec	Specifier
T	Tense
TOP	Topic marker

Acronyms

ASL	American Sign Language
CDI	Communicative Development Inventory
CLLD	Clitic Left Dislocation
CHILDES	The Child Language Data Exchange System
DLD	Developmental Language Disorder
DOR	Direct Object Restriction
EM	External Merge
ERP	Event-related potential
FL	Faculty of Language
fNIRS	Functional near-infrared spectroscopy
FOFC	Final-Over-Final Constraint/Condition
GLMMs	Generalized linear mixed-effects models
HP	Head Parameter
IM	Internal Merge
IPLP	Intermodal Preferential Looking Paradigm
LCA	Linear Correspondence Axiom
LMMs	Linear mixed-effects models
MLU	Mean Length of Utterance
NP	Nominal Phrase
PF	Phonetic Form
PLD	Primary Linguistic Data
RoIs	Regions of Interest
SD	Standard Deviation
SE	Standard Error mean
UG	Universal Grammar
UTAH	Uniformity of Theta Assignment Hypothesis
VEPS	Very Early Parameter Setting
VP	Verbal Phrase
vP	Little v
WALS	The World Atlas of Language Structures
WS	Workspace
WWO	Weird Word Order Paradigm

List of Tables

List of Figures

1 Acquisition of early grammar

The sound repertoires of the more than 7000 languages that are still around today (Eberhard, Simons, & Fennig, 2021) vary widely, as do their grammatical structures, and the meanings that their lexical items encode. However, despite these differences, children master their mother tongue, either spoken or signed, within about the first five o six years of life and with little or no explicit instruction. In fact, ignoring any attempt of instruction (McNeill, 1966). They command their native language at a remarkable level of complexity and computational sophistication well before they are capable of lacing their shoes or performing even simple mathematical operations such as addition or subtraction.

This chapter explores the theoretical approaches about the developmental origins of word order, asking by what mechanisms infants first begin to build a syntax. Syntactic bootstrapping (Gleitman, 1990; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005) begins with an unlearned bias determined by UG toward one-to-one mapping from each NP in a sentence to a theta-role in the event construal. Since UG defines possible human languages, its principles determine children's initial hypotheses and predict what will not occur in child (and adult) languages. This approach is subsumed by innatism following the idea proposed by Chomsky (1959) that children are equipped with an innate template or blueprint for language, which is called the Language Acquisition Device (LAD)(see Guasti, 2002 for more details).

In contrast, in lexical approaches (Abbot-Smith & Tomasello, 2006; Tomasello, 2003), or Piagetian constructivism in general (see Piatelli-Palmarini, 1980 for discussion), children begin with rote-learned concrete holophrases or low-level, lexically specific slot-and-frame patterns.

Only gradually do children abstract across these holophrases and lexical patterns to arrive at fully abstract, adult-like syntactic representations. Such approaches focus on words or word-like constructions (e.g. 'chunks'), do not acknowledge the innate mechanisms, and view acquisition essentially as statistical (Bates et al., 1996; Dąbrowska, 2015). It is worth pointing out that although Yang (2002, 2004)'s variational model (see Chapter (3.1)) places emphasis on statistical and probabilistic learning mechanisms, it still recognises the role of UG, i.e. a probabilistic/variational learning algorithm should operate in the domain-specific space of syntactic parameters.

The two broad approaches differ in their answers to the question, first, do infants have innate knowledge of linguistic categories and phrases (e.g. [NOUN], [VERB], [VP], [NP], [AGENT], [PATIENT]? Second, do infants necessarily go through different steps to arrive at adult-like abstract constructions? If the answer for the first question is no, and for the second is yes, then you are adopting lexical approaches, otherwise you are adopting syntactic bootstrapping. Thus, the purpose of this section is to review the main claims of each approach and the empirical evidence adduced for each of them. The debate between these two views lies the back of the theoretical approach of the thesis.

1.1 Poverty of stimulus

Results from corpus studies show that, already in the two-word stage, English-speaking children produce the VO order (as in (1a), example taken from Wells's corpus (1981) in CHILDES, MacWhinney, 2000) and their Japanese peers OV order (as in (1b), example taken from Yokoyama and Miyata's corpus (2017) in CHILDES).

(1) a. I cut it. (Abigail, 1;8)

b. Mikan tabe-yoo. (Kiichan, 1;8)
Orange eat-ORT
'Let's eat orange.'

The fact seems to indicate that children know what sort of language they are being exposed to[1], at least by the time they start putting words together, at roughly two years of age.

However, constituents are often moved. Although the canonical structure in Mandarin is SVO, the object can be moved to preverbal position, leading to a non-canonical, but grammatical SOV structure with a morphological marker *ba* (as in (2), example taken from Zhou's corpus (2001) in CHILDES).

[1]Notice that Japanese children around 2;5 sometimes produce VO order, like in (i), taken from Miyata's corpus (2004) in CHILDES.

(i) Yomoo, koko. (Aki, 2;7)
read-ORT this part
'Let's read this part.'

However, in fact, English-like SVO order is also available in Japanese, as illustrated in (ii), taken from Sugisaki (2008), yet exhibits various syntactic restrictions.

(ii) Eri-ga tabeta yo, sushi-o.
Eri-NOM ate EXL sushi-ACC
'Eri ate sushi.'

Sugisaki (2008) showed that the production of VO sentences in child Japanese has the same syntactic restrictions as for adults, as they are never used to formulate direct object *wh* questions, besides the frequency of occurrence of VO sentences is not different from that of adults.

(2) Ta ba ping-guo reng diao le. (Marui, 2;8)
Ta BA apple throw away PERF
'He threw away the apple.'

The noun phrase (NP) *ping-guo* 'apple' bears the same thematic relation to the verb *reng* 'throw' in both canonical SVO (see (3)) and non-canonical S*ba*OV, but appears preverbally in the latter.

(3) Ta reng diao le ping-guo.
Ta throw away PERF apple
'He threw away the apple.'

On the assumption that thematic relations are established in a strictly local fashion, constructions like (2) entail that the NP is displaced from its original position, i.e. the object is moved to the post-*ba* position. Imagine that our child, Marui, at 2;8 does not have any prior knowledge of phrase structure. He could formulate a rule like (4) on the basis of the input.

(4) Move the object to the position right after *ba*.

However, Marui would run into trouble as soon as he encountered sentences with various objects like in (5):

(5) Mama xi le Marui chuan de yi-fu.
mom wash PERF Marui wear DE clothes
'Mom washed Marui's clothes.'

Thus, if he applies the rule in (4) and moves the first object to the position right after the *ba*, he will get an ungrammatical sentence like (6a); or if he moves the second object, he will get the ungrammatical sentence as (6b) or (6c)

(6) a. * Mama ba Marui xi le chuan de yi-fu.
Mom BA Marui wash PERF wear DE clothes
'Mom washed Marui's clothes.'

b. * Mama ba yi-fu xi le Marui chuan de.
Mom BA clothes wash PERF Marui wear DE
'Mom washed Marui's clothes.'

c. * Mama ba chuan de yi-fu xi le Marui.
Mom BA wear DE clothes wash PERF Marui
'Mom washed Marui's clothes.'

However, to our knowledge, infants never utter sentences like (6). So the rule (4) is out, infants seem to operate differently. There would be an infinite number of possible hypotheses to consider, therefore there can never be enough input data to rule them out. Such a fact let Chomsky (1964, 1980) to formulate "the argument from poverty of the stimulus". As Chomsky (1986, p. XXV, from Russell, 1948) states:

> .... How comes it that human beings, whose contacts with the world are brief and personal and limited, are nevertheless able to know as much as they do know?

He identified the problem, referring back to Plato and Descartes, as humans possess innate knowledge, that is, knowledge which cannot have been derived from the environment "since the stimulus [to that knowledge] does not resemble what the mind produces on the occasion of stimulation" (1980, p. 35). He argued this knowledge to be language-specific and to constitute was is known as Universal Grammar (UG), which is the initial stage of the Faculty of Language (FL).[2] FL is a sys-

[2]One should not confuse UG with "linguistic universals", as the one

tem of natural computation plus an associated array of interfaces with other cognitive systems (these are called interface systems, and they include the sensory-motor and conceptual-intentional systems), which are understood to be one of the components of the human mind/brain.

Evidence compatible with innatism comes from both sign languages, spoken languages, brain-imaging studies and studies in Developmental Language Disorder (DLD). Some of the evidence collected over the years will be summarized next.

First, in the so-called "home sign", that is, when deaf children who are not exposed to conventional languages invent gestural communication systems that bear striking similarities to conventional sign languages (Goldin-Meadow, 2003). The creation of Nicaraguan Sign Language suggests the existence of some innate tendency to regularization despite the fact that the environment is markedly impoverished (Senghas, 1995; Senghas & Coppola, 2001; Senghas, Coppola, Newport, & Supalla, 1997). Children learning American Sign Language (ASL) from parents who are late learners of sign language look like other native speakers, that is, they learn ASL without the learning errors of their parents (Singleton & Newport, 2004). In hearing communities, this might arise from input they receive from native speakers outside the family, however, in the deaf community, many children learn ASL only from their late-learning parents, without exposure to native signers (Newport, 2020). For example, home signers commonly use or-

like Greenberg (1963). Linguistic universals are generalizations based on surface phenomena, or what Chomsky (1986) called E-languages, which with no doubt can provide valuable data about human language, but they often have exceptions.

der in a gesture string to mark abstract AGENT and PATIENT roles of diverse events (Coppola & Newport, 2005). Hunsicker and Goldin-Meadow (2012) provided evidence that homesigners string sequences like demonstrative and noun together as [[[that] [car]] [press]], rather than just producing linear flat sequences of the three signs [[that] [car] [press]], thus there must be a strong internal pressure to make linguistic expressions hierarchical even in the absence of experience (as opposed to Koko, a gorilla who has been taught more than 1000 signs based on ASL, but lacks syntax and grammar (Patterson, 1978; Patterson & Cohn, 1990); see also the case of Nim Chimpsky, an infant chimpanzee (Terrace, Petitto, Sanders, & Bever, 1979))[3].

In spoken languages, children say things they have never heard (see example in (7) from Marcus & Davis, 2019, emphasis mine), and resist any explicit corrections (see example in (8) from McNeill, 1966).

(7) Alexander: What's **chest-deep** water?

Mama: Chest-deep water is water that comes up to your chest.

Papa: It's different for each person. Chest-deep for me is higher than it would be for you.

Alexander: Chest-deep for you is **head-deep** for me.

[3]Actually, as has been pointed out recently by Petitto (2005), chimps do not actually learn words. For Nim, for instance, an "apple" was the object associated with the place that apples were kept, or the knife used to cut it, or the person who last gave him an apple...and so on and so forth, that is, only a hodge-podge of associations without any rules, which clearly differs from young human infants (see also Yang's test (2013) based on the data from Terrace et al., 1979).

(8) Child: Nobody don't like me.

Adult: No, say 'nobody likes me'.

Child: Nobody don't like me. (8 repetitions of this dialogue)

Adult No, now listen carefully; say 'nobody likes me'.

Child: Oh! Nobody don't likes me.

Second, language acquisition is subject to critical period effects as expected if it is innate. Critical periods are observed in various domains. In nonhuman animals, for example, a chaffinch will never learn the song correctly if it does not hear an adult song before it reaches sexual maturity (Thorpe, 1958)[4]. In humans, Banks, Aslin, and Letson (1975) show the existence of a critical period for susceptibility of binocular vision, roughly between 1 and 3 years of age and the absence of normal binocular experience during this period will cause irreversible changes or occasionally total lack of binocular functions (see also Mendola, Conner, Roy, & Chan, 2005 on experience-dependent neuronal plasticity in amblyopia). Lenneberg (1967) stated that language can develop fully only if it acquired before puberty. The case of Genie (Curtiss, 1977) shows that even after several years of linguistic rehabilitation, the linguistic abilities, especially syntax, were still very limited, precisely because her lack of linguistic and social interaction before the age of 13. The evidence from the acquisition of ASL at different ages also indi-

[4]With some clear similarities between songbird and human language, it has been proved that birdsong is only a model for speech, rather than language, as there is no hierarchical structure, structure dependency, nor the apparent displacement in birdsong (Berwick, Beckers, Okanoya, & Bolhuis, 2012; Berwick & Chomsky, 2017).

cates that its production and comprehension decline linearly with age of first exposure (Mayberry & Eichen, 1991). More recently, Hartshorne, Tenenbaum, & Pinker (2018) showed the existence of a sharply-defined critical period even for second language acquisition by analysing the data from 2/3 million native and non-native English speakers, suggesting that L2 learners cannot become native-like, at least in some domains after some certain ages.

Third, Mahmoudzadeh et al. (2013) using functional near-infrared spectroscopy (fNIRS) tested 28 weeks of premature infants' neural response to linguistic and non-linguistic stimuli. Their results showed infants at this age can discriminate linguistic sounds from non-linguistic voice, even when the cortical organization in layers is not completed and neural connectivity and neural migration are still developing, showing that infants are born with an innate bias to attend to human language. Musso et al.'s (2003) research, showed that the processing of expressions observing UG principles increased the activity of Broca's area, as expected, but that this does not hold for comparable expressions formed or interpreted by reliance on linear order—say, putting the negation word *no* always after the third word of the sentence or a *yes/no* question being formed by inverting the order of the single words —though that is a much simpler computation than the actual ones that are used (see also the case of Christopher, Smith & Tsimpli, 1995).

Finally, systematic comparisons of children with and without DLD reveal that the incidence of DLD is about 22% among members of the family of a child with DLD, and only 7% among members of the family of a child without DLD (Rice, Haney, & Wexler, 1998). This family aggregation suggests that DLD has a genetic basis. Although there is no evidence for a link between a specific gene and language disorders

(Chabris et al., 2013), several genes have been reported to associated with DLD (see Rice, 2012 for a review).

1.2 Grammatical approaches

Infants can get verb interpretation only through the syntactic structure of sentences in which each verb appears before they know anything about the verb's semantic content. This process is called syntactic bootstrapping (Gleitman, 1990). The rationale behind it is that, if sentence structures are projected from their semantics, then to some extent the semantics itself may be recoverable from the syntax.

A transitive sentence, such as (9a), has exactly two core NP dependents, or arguments; while an intransitive sentence, such as (9b), has one argument.

(9) a. John pushes Mary.

b. John swam.

Guided by the Theta Criterion (Chomsky, 1981), the learner hears a sentence with a novel word, like *tickle* in *John tickles Mary*, she will then understand that this sentence describes an event that is transitive. Then if there are just some events that are salient to her in the context, then she can restrict what *tickle* might mean. Besides, since we suppose that infants have abstract representations of linguistic experience, this predicts rapid transfer of syntactic knowledge to new verbs. For *John tickles Mary*, an early abstraction learner might represent the positions of nouns in sentences with *tickle* as preverbal and postverbal nouns, rather than only pre- and post-*tickle*.

In addition, their abstract semantic representations might indicate that John is an AGENT, and Mary a PATIENT by resorting to the Uniformity of Theta Assignment Hypothesis (UTAH, Baker, 1988), according to the AGENT asymmetrically c-commands the PATIENT (or THEME), thus are higher than PATIENT in the syntactic structure; or the Canonical Alignment Hypothesis (Hyams, Ntelitheos, & Manorohanta, 2006), according to which external arguments (in particular AGENTS) will map onto subject position. Given these representations, infants know from the start where AGENTS and PATIENTS belong in sentences; even upon encountering a new verb, they can rapidly use this knowledge to parse the sentence.

1.2.1 Acquisition evidence for syntactic bootstrapping

Regarding the evidence for the syntactic bootstrapping hypothesis, most work has tested syntactic bootstrapping in infants between 17 and 40 months. Starting with Naigles (1990), the Intermodal Preferential Looking Paradigm (IPLP, Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987) allows us to study the comprehension of sentences by infants. Naigles (1990) tested 2-year-old English infants' comprehension of both transitive and non-transitive actions with the pseudo-verb *gorp*. In the training phrase, half of the participants heard a transitive sentence as in (10a) and half an intransitive sentence as in (10b).

(10) a. The duck is gorping the bunny.

b. The duck and the bunny are gorping.

Both groups watched a scene in which a duck performed an action on a bunny on one screen and, on the other, the duck and bunny each performed a synchronous non-transitive, reflexive action. In the test phase, infants were asked to *find gorping*. Naigles (1990) found that infants who had heard the transitive frame looked longer at the transitive scene than those who heard the intransitive sentences, while children who heard the conjoined-subject intransitive audio displayed the opposite tendency (see also Naigles & Kako, 1993; Naigles, 1996 using slightly different methodology). Hirsh-Pasek and Golinkoff (1996) extended the results to 17-month-old infants and found that they can use word order to understand transitive sentences containing familiar words. Infants in Hirsh-Pasek and Golinkoff's study (1996) saw Big Bird tickling Cookie Monster on one screen and Cookie Monster tickling Big Bird on an adjacent screen, and then they were asked to look at where Cookie Monster was tickling Big Bird after hearing the sentence as in (11).

(11) Cookie Monster is tickling Big Bird.

Results show that 17-month-olds can use word order information to infer which scene corresponds to the sentence they have heard. One may think that the child cannot tell the meaning of *gorping* or *tickling* by only looking at the syntax, but from the syntax and the scene together, the meaning of the novel verbs is determined. That is what many *chase/flee* studies show (Gleitman, January, Nappa, & Trueswell, 2007; Nappa, Wessell, McEldoon, Gleitman, & Trueswell, 2009), in the sense that those verb pairs pose a special problem for observation, namely, one cannot tell *chase* from *flee* by looking at a scene as their situational contexts overlap very closely, but from the syntax and the scene together, their meaning is determined.

Surprisingly, Yuan and Fisher (2009) showed that toddlers can gather information of unknown verbs only from listening experience. They showed 28-month-olds dialogues in which they heard a novel verb *blick* in transitive (e.g., *Jane blicked the baby!)* or intransitive sentences (e.g., *Jane blicked!).* Then they viewed a test event while hearing the novel verb in isolation (*Find blicking!*). Children who had heard transitive dialogues looked longer at the two-participant event than did those who had heard intransitive dialogues. This effect held even when children were tested the next day, but disappeared if children heard only a neutral audio (i.e. *What's happening?*) during the test phrase. Messenger, Yuan, & Fisher (2015) extended the results to 22-month-olds, they showed that 22-month-olds can correctly interpret *Bill was blicking the duck* as a causative two-participant event, as opposed to a one-participant event, such as *Bill was blicking* even without visual cues in the learning phase. So, despite the uncertainty about the semantics of the verb, infants still can learn something about the verb (such as its transitivity) through its syntactic environment.

Another study, conducted by Gertner, Fisher, & Eisengart (2006) found that 21-month-old English-speaking children can use canonical word order to interpret transitive sentences containing pseudo-verbs. Test sentences were illustrated by two simultaneous videos with theta-role reversal as in the Hirsh-Pasek and Golinkoff's study (1996): one representing the target SVO interpretation and the other the non-target OVS interpretation. Yuan, Fisher, & Snedeker (2012) extended the results to 19-month-olds. They showed two video events, one was a two-participant event accompanied by a pseudo-verb in transitive sentences (e.g. *She's gorping her*), another was a one-participant event with intransitive sentences (e.g. *She's gorping*). They found that infants who heard transitive sentences looked longer at the two-participant event than did those who heard intransitive sentences. Moreover, Jin and Fisher (2014) adapted Yuan et al. (2012)'s experiment for even younger infants, 15-month-olds, with the same result.

However, as pointed out by Franck, Millotte, Posada, & Rizzi (2013), the infants' preference for the SVO over the OVS order in Gertner et al. (2006) may reflect a universal AGENT-first bias (as postulated by Bates & MacWhinney, 1982; Bever, 1970, and, more recently, Lidz, Gleitman, & Gleitman, 2001) rather than the acquisition of grammatical properties of word order. In addition, the pairs of videos used in Yuan et al. (2012) and Jin and Fisher (2014) differed in the number of participants, two participants in the case of transitive event (e.g. one actor rotated another on a swivel chair) or one participant in the case of intransitive event (e.g. one actor bounced on a yoga ball). Thus the two videos contrasted not only regarding word order, but also with respect to the action *per se*, which may lead infants to focus on guessing the lexical meaning of the pseudo-verb, rather than just word order.

1.3 Usage-based approaches

Usage-based, or lexical approaches (e.g., Abbot-Smith & Tomasello, 2006; Tomasello, 2003), which may date back to at least Braine (1963), claim that infants' earliest linguistic representations are not adult-like and grammatical development is acquired via the lexicon, which is to say that children initially encode item-based constructions which express semantic or thematic relations, but lack structural relations. As stated by Tomasello (2000, p. 210): "young children's creativity - productivity - with language has been grossly overestimated; beginning language learners produce novel utterances in **only some fairly lim-**

ited ways" (emphasis mine). On this view, very young children's language shows only a limited degree of combinatorial diversity, which bears similarities to imitation.

Basically, what lexical approaches propose is that at the early stage, young infants start at the same point as Nim Chimpsky (Terrace et al., 1979), a chimpanzee who only learned sign language based on storage and retrieval of item-based expressions, without abstract syntactic knowledge. The only difference between Nim and a human infant is that the later possesses a more general cognitive mechanism based on the input which is not domain-specific to language learning, but more general (e.g. inferential mechanisms). So a large amount of input is crucial in lexical approaches. In this framework, there are no functional categories, because inflection is stored in inflected words, in the sense that words and sentences (including idiosyncratic patterns) are all constructions. That is, there is no clear-cut point between the lexicon and the phrase-level component. Since the learning of early grammar is critically dependent on the child having a substantial store of relevant lexical items over which generalization could be made, abstract syntactic knowledge will not emerge until age 3 or 4 (Matthews, Lieven, Theakston, & Tomasello, 2007; Pine & Lieven, 1997; Tomasello, 2003)[5]. Although lexical approaches have often been interpreted - by both their critics and their advocates - as claiming that "children do not have abstract syntactic knowledge until age X" (cf. Franck et al., 2013; Savage, Lieven, Theakston, & Tomasello, 2003, 2006), it's fair enough to say,

[5]Although Elena Lieven (p.c.) agreed that their new studies also lead to the conclusion that even 21 months infants can understand the propositional content, as scopal semantics.

as has been pointed out by Ambridge and Lieven (2015), that the central claim of lexical approaches is not to say that until some relatively advanced age (say, age three or four) all or most of the children's knowledge consists of item-based structures, rather children will show better linguistic performance (in whatever task) when they can make use of string or lexical schema that they have frequently encountered in child-directed speech, and worse performance on an utterance for which no stored string is available. Besides, even when children have formed adult-like abstract grammatical knowledge, they will still show better performance for utterances that constitute prototypical representations of those constructions. It is unclear whether adults would do the same.

Along this line, infants will start out with rote-learned concrete holophrases (like Nim) without abstract syntactic knowledge. Some lexically specific slot-and-frame patterns or schemas that infants start with can be found in (12):

(12) a. I + want + it.
b. I'm [X]ing it.
c. It's a [X].

Pine and Lieven (1993) argued that 77% of children's naturalistic speech could have been generated by one of just 10 schemas as those in (12). Lieven, Salomo, & Tomasello (2009) found that 58%-78% of children's utterances were either repetitions of previous utterances or could have been generated by a single operation. Bannard, Lieven, & Tomasello (2009)'s mathematical model can generate 60-80% of children's utterances at 2;0 by using only slot-and-frame patterns like in (12). However, the issue remains as to how to account for the 40-20% of utterances which do not follow that pattern, and which are unexpected.

Moreover, according to Dodson and Tomasello (1998), for younger children, the 90% of all arguments in the transitive patterns in English were pronouns (e.g. *He is tamming it*), suggesting that they were relying heavily on the use of the slot-and-frame pattern such as (12b). Investigating the use of pronouns more directly, Childers and Tomasello (2001) found that two-year-olds showed significantly better performance when asked to produce novel transitive sentences with pronouns (e.g. *He is meeking it*) than with full lexical NPs (e.g. *The dog is meeking the car*)(85% vs. 45%).

Regarding word order acquisition, lexical approaches predict that infants understand the meaning of the verb first and memorize fragments they hear, and store in memory individual items with the associated syntactic environments. For a sentence like (13), they should know the meaning of the verb *push* first.

(13)	John	pushes	Mary.
	[pusher]	[the verb PUSH]	[pushee]

Once they know that, they know that it requires two roles: one is the pusher and another is someone who is being pushed[6]; then according to the (visual) situation or context, they associate the first NP as the AGENT and the second NP as the PATIENT. After having acquired an inventory of item-based constructions like *John kisses Mary*, *John hits Mary*, etc. from the input, they know that in English the

[6]It is alike to Dowty's proto-role view (1991) in the sense that there are different sets of role (as 'pusher' and 'pushee'; 'giver' and 'givee', etc.) specific to each verbal lexical item in a language, and that is what a child has to memorize.

AGENT corresponds to the subject position, and the PATIENT to the object position, thus at that stage they arrive at an abstract [AGENT] [ACTION][PATIENT] schema or a fully abstract syntactic SVO representation by analogical generalization. This makes argument structure a memorized list. Note that the whole process is gradual, and requires many input sentences, so frequency plays a crucial role here.

For non-canonical constructions (although it should not make a different whether it is canonical or not, for lexical approaches), they suggest that infants also start out with item-based constructions (e.g. *It's broken.*), and then they abstract it to form, first, lexically-specific schemas (e.g.*[X] is broken.*) and, ultimately, abstract syntactic construction (e.g.*[X] BE/GET [Y] by [Z]* (Savage et al., 2003). Since '*[X] BE/GET [Y] by [Z]*' is a non-canonical passive in English, which is rare in spoken discourse, children have difficulty with this structure even at age 6 (Messenger, Branigan, & McLean, 2012), and some adult speakers may never acquire a fully abstract passive construction (Ambridge & Lieven, 2015; Ambridge, Bidgood, Pine, Rowland, & Freudenthal, 2016; Dąbrowska & Street, 2006). So for non-canonical constructions, lexical approaches would predict that infants would have worse performance compared to the canonical ones, since the frequency of the first is very low in the input.

In summary, on this account, there is no abstract knowledge initially and long-term exposure is required and only at later stages does the child generalize through analogy from memorized fragments to abstract notions such as general word order properties. Therefore, lexical approaches contrast with syntactic bootstrapping, according to which children learn form and meaning about each word, but at the same time are also biased to represent syntactic and semantic facts of sentences in abstract terms.

1.3.1 Acquisition evidence for and against lexical approaches

Akhtar (1999) was the first to use the Weird Word Order (WWO) paradigm and argue in favour of the lexical model on the basis of her results. She found that English-speaking two-year-old and three-year-old children were more likely to produce the SOV ('weird') word order rather a canonical SVO construction after hearing a non-canonical SOV sentence with a novel verb (e.g., *Elmo the car meeking*) than when hearing a familiar verb in a SOV sentence (e.g., *Elmo the car pushing*), in which case they tended to correct the sentence. However, the four-year-old children were more likely to correct the SOV construction, and to explain the event using a reorganised canonical SVO word order even in the pseudo-verb condition. Adapting Akhtar's (1999) methodology, Matthews et al. (2005) used real English verbs, and let children (aged 2;9 and 3;9) hear either high or low frequency verbs modelled in ungrammatical SOV order. They found that children aged 2;9 who heard low frequency verbs were significantly more likely to adopt the WWO than those who heard higher frequency verbs. Children aged 3;9 preferred to use the grammatical SVO order regardless of verb frequency. The same evidence can be found in French (see Matthews et al., 2007): French children (aged 2;10 and 3;9) heard either ungrammatical SOV or ungrammatical VSO order with high or low frequency verbs. Results showed that children were more likely to adopt a WWO if they heard lower frequency verbs and children in the high frequency conditions tended to correct the ungrammatical orders to the grammatical alternative.

The fact that younger children tended to reuse the weird word orders modelled by the experimenter when the sentence contained low-frequency verbs in the previous studies was taken to support the lexical approaches according to which it is not until some time after age 3;0 that children develop abstract knowledge of word order. In a similar vein, the syntactic priming study of Savage et al. (2003; see also 2006) concludes that abstract priming does not occur until after age 4;0, a finding that has been contested by several authors (see counter-evidence in Bencini & Valian, 2008; Rowland, Chang, Ambridge, Pine, & Lieven, 2012; Shimpi, Gámez, Huttenlocher, & Vasilyeva, 2007).

A recent study by Abbot-Smith, Lieven, & Tomasello (2008) found that when presenting 2-year-old English-speaking children with ungrammatical OVS sentences with pseudo-verbs (rather than SOV or OSV as in previous studies), children corrected them more often. This fact leads the authors to propose that by age 2 children have some weak and gradually abstract grammatical knowledge. Besides, cross-linguistic variation also emerges: German children correct WWO more often than English children with ungrammatical OVS sentences. The authors argue that this difference is due to the presence of case marking in German, which is a strong cue for thematic role assignment (Abbot-Smith et al., 2008).

Besides, in all of the studies using WWO paradigm, when correcting to SVO, children used pronouns frequently, but never when imitating SOV or VSO. According to the authors, this suggests that infants rely heavily on slot-and-frame patterns like *He's [X]ing it* detailed above.

The WWO methodology is often used to support the lexical approaches. However some critical issues with respect to this paradigm have been raised recently by Franck and Lassotta (2012) and Lassotta (2021). First, as pointed out by Franck and Lassotta (2012), it is hard to conceive the concept of abstractness as a gradable property. Second, it is incorrect to assume that if children have abstract syntactic knowledge, they will correct ungrammatical word order, as to correct ungrammatical word order, infants need to have the ability to infer the experimenter's expectation as well, namely, they need to know that the experimenter expects them to correct the incorrect structure from the input specially when they hear the experimenter speak normally during all the training phase. The method is particularly confusing when the children have been hearing the experimenter speak normally during the training phase, and then hear s/he turn to ungrammatical sequences. Franck, Millotte, & Lassotta (2011) replicated Matthews et al. (2007) with some slightly changes. They presented test sentences via a computer voice, rather than the experimenter, thus reduced the tendency of the child to imitate the experimenter and they obtained significantly more WWO corrections (see also Chang, Kobayashi, & Amano, 2009 for the same results in Japanese). Besides, Franck et al. (2011) also found that when grammatical sentences were introduced (rather than involving exclusively ungrammatical sentences as in Matthews et al., 2007), the younger group (aged 2;11) corrected ungrammatical sentences at a similar rate than the older group, as opposed to Matthews et al. (2007). Thus introducing grammatical models helped children realise that ungrammatical models were weird.

Third, after reanalyzing the data proposed in four studies on WWO (i.e. Akhtar, 1999; Abbot-Smith et al., 2001; Matthews et al., 2005, 2007), Franck and Lassotta (2012) found that actually young children seldom produce ungrammatical sentences (only 0.9% - 10% depending on the experiment), and older children re-use WWO at a similar rate to young children, despite the fact that they are assumed to have grammar. The authors reworked the recounts and found missing not-at-random data, young children re-used the ungrammatical word orders with novel verbs between 0.9% and 7.5% of their responses, while the rate for older children is between 0.3% and 16.3%. Moreover, the authors ran the same experiment on 8-9 year-old children, half of them systematically re-used the WWO presented by the experimenter (with effort), just because they believed that was what the experimenter asked for (Franck et al., 2011), given that 8-9 years olds have grammatical knowledge, this finding questions the WWO method altogether.

One important issue, as pointed out by Rizzi (p.c.), is that the absence of UG does not exclude *per se* the possibility that abstract grammatical knowledge is acquired early, since one can easily imagine that the use of domain general knowledge is smarter and more efficient than Tomasello (2003) assumes. Similarly, Wexler (1998) in his Very Early Parameter Setting hypothesis (VEPS) argues that parameters are set very early, but this does not deductively follow from any deep principle of Innatism: in fact various people sharing the innatist approach have proposed very different ideas on the time course of parameter setting (cf. Gibson & Wexler, 1994 vs. Yang, 2002, 2004, 2016 in chapter 3.1).

Thus this work touches on this debate. We can ask the question about which is the most suitable model of language acquisition that provides good fit to empirically data provided by the research here and in the literature. That is, do infants have abstract grammatical knowledge of word order? If they do, at which age?

1.4 Layout of the dissertation

With these general questions in mind, the layout of this dissertation is as follows. Chapter 2 contains a detailed description of the grammar of word order and the syntax of both canonical and non-canonical structures in Mandarin. In chapter 3, I conduct an analysis of a Mandarin Corpus from the CHILDES database (MacWhinney, 2000) with spontaneous production data by children from 1;2 to 6;5 and their caregivers to provide insight as to the role of input in the early development of word order and null arguments. While chapter 3 focuses on production, chapter 4 and chapter 5 explore infant comprehension of word order using eye-tracking techniques. I report two experimental studies examining both the acquisition of canonical (in chapter 4) and non-canonical (in chapter 5) word order representations of 17-month-old infants with pseudo-verbs. Chapter 6 gives a summary, draws conclusions and proposes future research.

2 Grammatical assumptions and research questions

In the minimalist program for linguistic theory (Chomsky, 1995, and subsequent work), there is only one fundamental operation of syntax, namely Merge, which takes two elements, α and β, and combines them into a set $\{\alpha, \beta\}$[7].

If α and β are two different syntactic objects, when merged into one structure, the kind of Merge is called External Merge (EM, Chomsky, 2004), and if the very same element, say α, is merged multiple times, the operation is called Internal Merge (IM, Chomsky, 2004). It is now standard to assume that Movement as we observed in (2), repeated here as (14) for convenience, is an instance of IM, thus both EM and IM are instances of the only one operation Merge, solely distinguishable when looked at from their source.

(14) Ta ba ping-guo reng diao le. (Marui, 2;8)
Ta BA apple throw away PERF
'He threw away the apple.'

Recently, Chomsky (2015, 2019) has claimed that IM is more complex than EM is 'a residue of the earlier errors' (Chomsky, 2015, p. 111).

[7] A set { }, by definition, imposes no intrinsic ordering among its members. Recently, Chomsky (2019) revises Merge to MERGE, an operation on workspace (WS) itself, rather than on particular syntactic objects, say P, Q. Then MERGE (P, Q, WS) = $[\{P, Q\}, X_1,\ldots,X_n]$ = WS'. In this sense, MERGE can be regarded as a mapping operation from one stage WS to another stage WS', but either way, Merge does not determine linear order.

If anything, IM is simpler, since IM keeps to a single syntactic object, while EM requires search of the entire workspace and lexicon. In addition, Chomsky (2004, 2015, 2019) also points out that the distinction between EM and IM appears at the conceptual-intentional interface: EM is related to argument structure ("who does what to whom", etc.), while IM correlates with discourse-oriented, information structures (as the scope of operators, topicality and focus, etc.).

2.1 Phrase structure

According to the X-bar template (Chomsky, 1970, 1986) indicated in (15), where for any category X (either lexical categories, as VERB, NOUN or functional categories like BA in Mandarin), there is a maximal projection that consists of the specifier (Spec) of X' and X' itself. In turn, X' consists of the head X and its complement (Comp) YP. All elements functioning in natural language syntax appear to exhibit the same three units of projection.

(15) XP $\rightarrow$ {Spec, X'}
X' $\rightarrow$ {X or X', (YP)}

Under the classical view, the order of the head and the complement may vary from language to language, leading to the VO order as in English and OV order as in Japanese, and from their base-generated positions, constituents can move to higher positions in the sentences.

In the minimalist program, since Merge is the only operation, bar-levels are considered relational properties, rather than categorical features. That is, in the classical X-bar Theory, X, X' and XP have different intrinsic features, while, in the minimalist approach to phrase structure, categories differ in virtue of their relations with elements in their local environment. In place of (15) we have only (16). Standard X-bar theory is thus largely eliminated in favour of bare essentials, i.e. Bare Phrase Structure (Chomsky, 1994, 1995).

In (16), the head α merges with β, then we label the resulting object as α, in another words, α projects. If it merges with another syntactic object, say γ, the whole structure K will be like (16), where the label of K is α.

(16) $\mathrm{K} = \{\alpha\ \{\ \gamma\ \{\alpha\ \{\alpha,\ \beta\}\}\}\}$

Notice that, if β is a lexical item, and does not project any more, then β is both a minimal and maximal projection. By this means, there are no vacuous projections or extra bar levels as in the X-bar Theory[8]. Besides, complements, specifiers and adjunctions are all maximal projections, since they do not project further.

In this way, linguistic structure is formed simply by reference to domination, and all syntactic variation is attributed to differences in the feature relations of particular items in the lexicon (following Baker, 2008; Borer, 1984; Chomsky, 1995) and/or at the syntax-sensorimotor interface (Chomsky, 2010).

[8]However, much of the discussion reported in this dissertation may use the traditional X-bar notation, familiar from textbooks, with phrasal nodes such as NP, VP, etc.

2.2 Word order parameters

For the three major constituents, subject (S), object (O), and verb (V), there are six (namely 3!) logically possible word orders. Most of the world's languages are reported to have a dominant or so called canonical word order (1188 out of 1377 in the WALS sample, see table 1).

Table 1: Word orders across world's languages (Dryer, 2013).

Word order	Number of languages
1. SOV	565
2. SVO	488
3. VSO	95
4. VOS	25
5. OVS	11
6. OSV	4
No dominant order	189

But these orders differ substantially in their frequency. SOV and SVO are most prevalent, while the third most common word order, VSO, just occurs in 8% of languages and the remaining three orders are extraordinarily rare. As originally observed by Greenberg (1963), most languages are either consistently head initial (prepositional, VO, Aux V, C initial) or consistently head final (postpositional, OV, V Aux, C final), although a minority of languages (German, Dutch) can be mixed.

Traditionally, word order has been viewed as a syntactic composition within UG (Chomsky, 1986), which is associated with the head parameter (HP, or directionality parameter Stowell, 1981; Travis, 1984). In head-initial languages like English, the head precedes its complement. Therefore, the verb precedes the object like in *The girl pushes* *the boy*, the preposition *on* precedes the DP *the chair*, and so on across all types of constituents. Head-final languages like Japanese or Turkish show the mirror order, with the head following its complement. Here, the object precedes the verb, the noun precedes the preposition, and so on. Unfortunately, the harmony proposed by Greenberg (1963) as mentioned earlier is empirically challenged. Mandarin is considered as a head-initial language, however, it exhibits some word order patterns that are unattested in other head-initial languages (see (17), Dryer, 1992, exemplified in (18)- (20)):

(17) The order of (subject and object) relative clause and noun: REL-N
The order of prepositional phrase and verb: PP-V[9]
The order of standard noun and adjective in comparative: ST-ADJ

(18) a. na ge mai shu de nan-hai
DET CLASS buy book DE boy
'the boy that bought books'

b. na ben Zhang-san xi-huan de shu
DET CLASS Zhang-san like DE book
'the book that Zhang-san likes'

(19) wo-shi li
bedroom in
'in the bedroom'

(20) Zhang-san bi Li-si gao.
Zhangsan BI Li-si tall
'Zhangsan is taller than Lisi.'

[9] Huang, Li, & Li (2009) argue that the apparent postpositions in Chinese actually behave more like nominal clitics, so they claim that the prepositional phrase in Chinese is still head-initial, i.e. P + NP.

The fact described above have made it impossible to describe language variation systematically by specifying languages as head-initial or head-final. However, the HP *per se* does not exclude the possibility of disharmony. Notice that the HP is based on X-bar template, in principle, the order of Complement and Specifier with respect to the head can be either on the left or on the right, even within the same language. For example, Hungarian admits both propositions and postpositions (Marácz, 1989). It seems that this disharmony with respect to word order can be captured by the HP. Nevertheless, precisely because of it, as also pointed out by Kayne (1994), a problem of the HP approach is that first, it will overgenerate structures that are not found in natural languages, especially we do not find languages which show the pattern with the Specifier to the right of the head, as for instance, *wh*-movement will not occur to the right of [Spec, CP]. One may claim that this was already assumed when there was a directionality parameter: the Spec was always to the left, no variation there.

Second, more importantly, it seems that the HP has nothing to say about a robustly attested correlation between word order and syntactic phenomena. Kayne (1994) advocated a single universal word order template: Linear Correspondence Axiom (LCA):

(21) The Linear Correspondence Axiom:
If α asymmetrically c-commands β, then α precedes β in linear order.

LCA maps from hierarchical structure to linear order. The empirical consequence of LCA is that there will be only one word order (SVO), since the specifier of a head will always asymmetrically c-command the head, in linear order, the specifier will always precede the head, thus rules out the Spec to the right.

Besides, according to the LCA, English or Mandarin VO and Japanese OV must differ in more than just order: the V in English must asymmetrically c-command O, whereas in Japanese that cannot be the case. If interpretation is not affected by the difference between VO and OV (Chomsky, Gallego, & Ott, 2019), and by assuming something like Baker's Uniformity of Theta Assignment Hypothesis (1988), then English VO and Japanese OV cannot be attributed to a difference in MERGE (external merge in this case), namely, it cannot be the case that English merges O as a complement of V while Japanese merges O as a specifier of V.

In addition, if we adopt LCA, then, OV can never be associated with a structure in which O is in the complement position of V, namely, O must be moved to some Spec position above V in OV structure (Kayne, 2018). If we assume the object is the complement of the V, then the order SVO, namely Spec-Head-Compl is universal, and all other variations are due to movement (internal merge). One condition for the proposal of the LCA to be plausible is that one has to assume that every phrase has complex structures, because even in VO order O is not the sister of V.

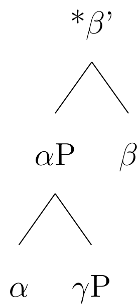
Beyond the LCA, there are some constraints about word order cross-linguistically, that is, certain orders are unattested in many syntactic domains, one of them is termed Final-Over-Final Constraint/Condition (FOFC) by Biberauer, Holmberg, & Roberts (2007, 2008, 2014):

(22) The Final-Over-Final Constraint/Condition (FOFC):
If α is a head-initial phrase and β is a phrase immediately

dominating α, then β must be head-initial. If α is a head-final phrase and β is a phrase immediately dominating α, then β can be head-initial or head-final.

Thus according to the FOFC, some orders as in (23) are not possible in natural languages.

(23)



Recently, ordering is considered as a reflex of the process of externalization of structures[10] by the sensory-motor system (Berwick & Chomsky, 2011). Chomsky (2013) further claims that linear order is only a peripheral part of human language, and it plays little role in the meaning-side of linguistic computation. Nevertheless in this thesis, I pursue the issue of the acquisition of word order assuming that it pertains to the acquisition of core grammar.

[10]By externalizarion it includes, according to Berwick & Chomsky (2017, p. 111) "morphophonology, phonetics, prosody, and all the rest involved in realizing spoken and gestured language or the parsing of speech or signed languages."

2.3 A sketch of the grammar of Mandarin Chinese

Here I consider some features of the grammar of Mandarin Chinese relevant for the development of the dissertation. I consider null arguments and word order in turn.

2.3.1 Null arguments in Mandarin

Mandarin permits the presence of null arguments, with both subject and object drop as long as reference can be recovered through the previous discourse context (Huang, 1984a). (24) is an example of topic drop (with both subject and object drop).

(24) a. Speaker A: Ni-men dou kan guo tai-tan-ni-ke-hao ma?
you all see EXP Titanic SFP
'Have you seen *Titanic*?'

b. Speaker B: (Ø) Kan guo (Ø).
see EXP
'(We) have already seen (it).'

In (24b), both the subject *wo-men* 'I' and direct object *tai-tan-ni-ke-hao* 'Titanic' can be dropped, because they can be recovered through the discourse provided by (24a). Thus, subject and object drop, as in this case, can be considered instances of topic drop. According to Huang (1989) unlike null objects, not all null subjects in Chinese are null topics. Compare the following interpretations of (25) as (25a) and (25b):

(25) Zhang-san shuo hen xi-huan tai-tan-ni-ke-hao.
Zhang-san say very like Titanic

a. Zhang-san shuo hen xi-huan tai-tan-ni-ke-hao. (Null
Zhang-san$_i$ say [Top e$_j$] very like Titanic
Topic)

'Zhangsan$_i$ said that (he$_j$) liked Titanic.'

b. Zhang-san shuo hen xi-huan tai-tan-ni-ke-hao. (Null
Zhang-san$_i$ say [*pro* e$_i$] very like Titanic
pro)

'Zhangsan$_i$ said that (he$_i$) liked Titanic.'

Embedded null subjects may refer to a topic, namely some other person in the discourse, as illustrated in (25a) or may be A-bound by the matrix subject *Zhang-san*, as in (25b). Thus, the null subject in (25b) is regarded as a null pronominal (*pro*).

Recently, Barbosa (2019) based on a generalization by Tomioka (2003) proposed a unified approach to analyse both Mandarin-like discourse pro-drop languages and Spanish or Brazilian Portuguese-like consistent or partial null subject languages.

She claims that all these crosslinguistic variations can be captured by the fact of whether the language permits a minimally specified NP (nP) in argument position. A minimally specified NP is an NP that lacks a syntactically projecting restricting property. nP may combine with functional categories (number, Person, D, etc.), when such a rootless nP is merged under an overt D, then an overt pronoun is obtained. Languages that have a null D will allow null arguments (as Brazilian Portuguese, etc.), languages that have bare nominals in argument position will allow this nP in argument position, thus they will have null subjects and null objects (as Chinese, etc.).

2.3.2 Canonical word order in Mandarin

Word order and animacy have been considered the most reliable syntactic devices for sentence interpretation in Mandarin and Cantonese Chinese (Chang, 1992; Li, Bates, & MacWhinney, 1993; Miao, 1981), given the lack of morphological markers such as agreement, number, gender or case in these languages. The canonical word order in Mandarin is SVO (Li, 1990; Sun & Givon, 1985), illustrated in (26), though these matters have been the subject of decades of controversy (see Li & Thompson, 1975; Tai, 1973 for the claim that Mandarin is SOV and pospositional; and see Chu, 1979, Huang, 1978, Sun & Givon, 1985 for the claim that it is SVO and prepositional).

(26) Xiao-tu-zi zhua le xiao-ya-zi.
little-rabbit catch PERF little-duck
'The bunny caught the duckling.'

Besides, Mandarin is a topic-prominent language (Li & Thompson, 1976; Huang, 1982), which means that almost every argument is allowed in the topicalized position, either high or low, leading to non-canoncial word orders.

2.3.3 Non-canonical word order in Mandarin

Apart from canonical SVO word order, three other word orders, SOV, OSV, and VOS, are also possible in the spoken language. But those non-canonical word orders are semantically and pragmatically marked in special ways and require or allow special morphosyntactic markers, such as object marker *ba* in (27a) or passive *bei* or intonation marking topichood of the sentence object in (27b), or the so called 'afterthoughts' or 'dislocation focus construction' (Chao, 1968; Lu, 1980) in (27c). On the other hand, OVS and VSO are ill-formed in Mandarin (Li et al., 1993).

(27) a. Ta ba ping-guo chi le. (SOV)
he BA apple eat
'He ate the apple.'

b. Ping-guo, ta chi le. (OSV)
apple, he eat PERF
'As for the apple, he ate it.'

c. Chi le ping-guo ta. (VOS)
eat PERF apple he
'He ate the apple.'

In the case of SOV, when the shifted object is inanimate, Chinese permits bare SOV sentences like (28). However, as noted by Li and Thompson (1981), they are pragmatically restricted to situations where the speaker provides information counter to the expectation of the listener. Thus in order to make (28) felicitous, one has to imagine a situation in which eating the apple would be unexpected.

(28) Ta ping-guo chi le.
he apple eat PERF
'He ate the apple.'

On the other hand, if the shifted object is animate, then SOV obligatorily requires the morpho-syntactic marker *ba* (Van Bergen, 2006) as in (29).

(29) Xiao-tu-zi ba xiao-ya-zi zhua le.
little-rabbit BA little-duck catch PERF
'The bunny caught the duckling.'

The moved objects are marked with *ba*, and the resulting construction is referred to as the *ba* construction, which I discuss below.

2.3.3.1 The *ba* construction

The pragmatic difference between the canonical SVO construction and the *ba* construction is, following Tsao (1987), that the function of *ba* is to mark the following NP as a special topic and to bring into focus the result (see also Gao, 1997, and more recently, Kuo, 2015 for the similar approach). In this sense, the NP following *ba*, reffed to as post-*ba* NP, can be followed by Topic markers like *ya*, *a*, or *ne* (Li & Thompson, 1981), as in (30):

(30) Wo ba zhe ge bao ne fang zhuo-zi shang le.
I BA this CLASS bag TOP put table up PERF
'I have put this bag on the table.'

Moreover, this topic may be contrastive, as illustrated by the contrast in (31), with the post-*ba* NP followed by the contrastive topic marker *ne*.

(31) Wo ba zhe ge bao ne fang sha-fa shang le, ba na
I BA this CLASS bag TOP put sofa up PERF, BA that
ge bao ne fang yi-zi shang le.
CLASS bag TOP put chair up PERF
'I have put this bag on the sofa and that bag on the chair.'

The history of the *ba* construction is long and the structure has been analysed in different ways (for traditional analysis, see Wang, 1945, 1947 or Zou, 1995 for an extensive review). *Ba* (把) was a lexical verb, with the meaning of ‘take hold of’ or ‘grasp’ (Wang, 1957), however, through a process of grammaticalization, *ba* has lost standard verbal properties: it cannot take an aspectual marker after it (32a) nor can it serve as a simple answer to a polar question (32b) (see Zou, 1995 for an extensive review).

(32) a. *Xiao-tu-zi ba le xiao-ya-zi zhua (le).
little-rabbit BA PERF little-duck catch PERF
‘The bunny caught the duckling.’

b. *(mei/bu) ba.
(not) BA

Although there is no consensus on the nature of the *ba* construction, it is agreed in the literature that there are several restrictions concerning this structure (see Chao, 1968; Li, 2006). First, the NP following *ba* cannot be omitted (see (33)), even if Mandarin is an object drop language.

(33) *Xiao-tu-zi ba zhua le.
little-rabbit BA catch PERF

Second, Li (2006) proposes that the object of *ba* (i.e. the post-*ba* NP) must be animate and/or referential (with definiteness or specificity). So that sentence (34) with a non-specific object is ungrammatical.

(34) *Wo ba san-wu kuai dan-gao chi le.
I BA three-five CLASS cake eat PERF
‘I ate three to five cakes.’

There is a proposal according to which *ba* is a preposition (Chao, 1968; Lü, 1980), with *ba* and the post-*ba* NP forming a constituent modifying the following VP. However the preposition analysis is not quite satisfactory because if one assumes *ba* to be a preposition, then it is expected that the post-*ba* NP forms a constituent with *ba*, namely, the post-*ba* NP alone (NP without *ba*) and the VP should not form a unit. To the contrary, *ba* can take [NP VP] as its complement in Chinese, as shown by the coordination test in (35) taken from Huang et al. (2009, p. 166).

(35) Ta ba men xi hao, chuang-hu ca gan-jing le.
he BA door wash finish window wipe clean PERF
'He washed the door and wiped the windows clean.'

In (35) the object forms a constituent with the following VP, namely [men xi-hao] and [chuang-hu can-gan-jing], not with *ba*, therefore, *ba* is not a preposition.

More recently, Huang et al. (2009) noticed that a *ba* sentence always has an acceptable non-*ba* counterpart (see (36a) and (36b)), which suggests that none of the arguments in (36b) depends on *ba* for thematic assignment. They assume that *ba* itself does not assign a theta role, but Case (according to Li (1985, 1990), Case is assigned to the right in Mandarin). *Ba* assigns Case to the post-*ba* NP and since Case assignment follows an adjacency condition (Stowell, 1981), *ba* must be immediately followed by an NP (see also Li, 1985, 1990).

(36) a. Xiao-tu-zi ba xiao-ya-zi zhua le.
little-rabbit BA little-duck catch PERF
'The bunny caught the duckling.'

b. Xiao-tu-zi zhua le xiao-ya-zi.
little-rabbit catch PERF little-duck
'The bunny caught the duckling.'

Along this line, Tse (2020) establishes an analogy between Romance differential object marking (DOM), which also assigns case, and Chinese *ba* and he suggests that DOM could be implemented in different domains. In the case of Romance, DOM is associated with the nominal domain in which the nature of the object is the primary trigger; while in Chinese *ba* applies in the verbal domain in which the nature of the verb is the primary trigger.

The former observation is in line with the approaches that treat *ba* as the head of the projection. Huang et al. (2009) propose that *ba* is a functional head of its own phrase, namely BaP and higher than the *v*P. Sun (2019), in a recent analysis of the *ba* construction in the cartographic framework, proposes that *ba* is the head of a resultative VoiceP that selects for the lower, preverbal object, in line with the hierarchy of functional projections in Cinque (1999, 2006) for the clause (i.e. the former IP/TP) as given in (37), taken from Cinque (1999, p. 106).

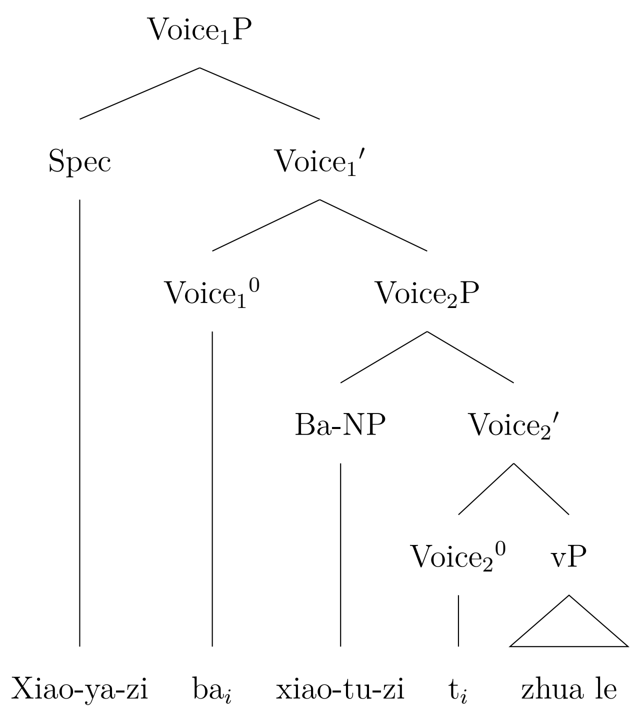
(37) $\text{Mood}_{\text{Speech Act}}$ $\text{Mood}_{\text{Evaluative}}$ $\text{Mood}_{\text{Evidential}}$ $\text{Mood}_{\text{Epistemic}}$ T(Past) T(Future) $\text{Mood}_{\text{Irrealis}}$ $\text{Mod}_{\text{Necessity}}$ $\text{Mod}_{\text{Possibility}}$ $\text{Asp}_{\text{Habitual}}$ $\text{Asp}_{\text{Repetitive(I)}}$ $\text{Asp}_{\text{Frequentative(I)}}$ $\text{Mod}_{\text{Volitional}}$ $\text{Asp}_{\text{Celerative(I)}}$ T(Anterior) $\text{Asp}_{\text{Terminative}}$ $\text{Asp}_{\text{Continuative}}$ $\text{Asp}_{\text{Perfect(?)}}$ $\text{Asp}_{\text{Retrospective}}$ $\text{Asp}_{\text{Proximative}}$ $\text{Asp}_{\text{Durative}}$ $\text{Asp}_{\text{Generic/progressive}}$ $\text{Asp}_{\text{Prospective}}$ $\text{Asp}_{\text{SgCompletive(I)}}$ $\text{Asp}_{\text{PlCompletive}}$ Voice $\text{Asp}_{\text{Celerative(II)}}$ $\text{Asp}_{\text{Repetitive(II)}}$ $\text{Asp}_{\text{Frequentative(II)}}$ $\text{Asp}_{\text{SgCompletive(II)}}$

According to Sun (2019), *ba* corresponds to a Voice head in the hierarchy because of its placement relative to other functional heads and adverbs like *hao-hao* ‘well’ (which corresponds to the Voice head), together with its semantic import as in (38).

(38) Ta cheng-gong de hao-hao ba dan-gao chi diao le.
he success ADV well BA cake eat finish PERF
‘He finished eating the cake successfully.’

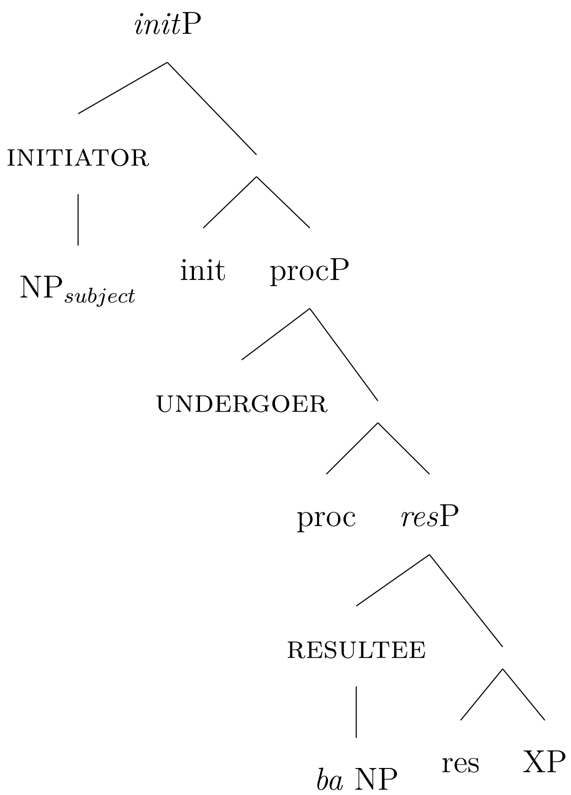
In addition, in order to capture the fact that nothing can intervene between *ba* and post-*ba* NP, Sun (2019) proposes that the VoiceP is split in two: the higher $Voice_1P$ holds *ba* in the head position, while the lower $Voice_2P$ holds a trace of the raised *ba* or a silent voice head to encode the active value. This head then attracts the object from the event structure to the specifier of $Voice_2P$. Thus, the structure of sentence (29) is in (39).

(39)



Sun’s analysis (2019) attempts to predict the restrictions on theta-role structure of the *ba* construction. She adopted event-decomposition structures proposed by Ramchand (2008) with three subevents (initiation, *process* and *result*), which are associated with the argument role of INITIATOR, UNDERGOER and RESULTEE respectively[11] to better explain the event properties of the *ba* construction in comparison to the traditional *v*P analyses. A maximal event-decomposition for the *ba* construction can be represented as the structure in (40):

(40)



In this structure INITIATOR hosts the external theta role, i.e. the subject of a *ba* construction; while the RESULTEE hosts the internal argument,

[11]INITIATOR is responsible for the eventuality coming into existence; UNDERGOER contains some sort of change or transition, but not necessarily attains a final state; and RESULTEE should attain an identifiable change of state (Ramchand, 2008).

i.e. the post-*ba* NP, this captures the complexity of the construction, which should be active and resultative at the same time (for the similar claims on the affectee character of the post-*ba* NP, see also Jheng, 2012; Tsao, 1987). This correctly excludes psychological verbs from entering the *ba* construction as in (41).

(41) *Xiao-tu-zi ba xiao-ya-zi heng le.
little-rabbit BA little-duck hate PERF
'The bunny hates the duckling.'

Besides, the INITIATOR is always distinct from the RESULTEE, such relation excludes all intransitive events entering the construction, and predicts that the object must be syntactically active[12]. The issue of the post-*ba* NP not being possibly null needs to be accounted for, although, to my knowledge, there is no discussion on that in the literature.

There are many other ways to analyse the *ba* construction (see for example, the aspectual analysis by Liu, 1997 or Sybesma, 1992), however, I will not delve into this issue any further here, since all approaches view *ba* as a functional head and suggest that the direct object moves from its canonical position to the post-*ba* position.

[12]The fact that the post-*ba* NP has to be overtly realized in Mandarin makes *ba* behave more like a preposition (Li, 1990). In this analysis, if *ba* is a preposition, then the post-ba NP is an NP, part of a PP. However, as I mentioned earlier, the prepositional analysis has been ruled out.

2.3.4 OS*ba*OV

As a topic-prominent language (Li & Thompson, 1976, 1981), the object of Mandarin can appear in a canonical Top(ic) position in the left periphery, together with a coindexed resumptive pronoun as post-*ba* NP in preverbal position as in (42).

(42) Xiao-tu-zi, xiao-ya-zi ba ta zhua le.
the-little-rabbit$_i$ little-duck BA it$_i$ catch PERF
'The bunny, the duckling caught it.'

Xiao-tu-zi 'the bunny' in (42) has been topicalized and appears in the sentence-initial position. Besides, this left-dislocated topic *xiao-tu-zi* 'the bunny' is coreferential with the overt clitic pronoun *ta* 'it' as an overt object in the *ba* construction is mandatory. In addition, prosodically there is a pause between the left-peripheral NP (i.e. the object) and the second NP (i.e. the subject).

Finally, note that the sentence-initial topic in Mandarin is an unmarked option, while the intra-clausal topic, as the *ba* construction is the marked one (Yu & Zhang, 2019).

To sum up, there are three ways to externalize the object in Mandarin, represented in (43).

(43) a. Object in VP: SVO
b. Object in vP or above, but below CP: S*ba*OV
c. Object in the left periphery: OS*ba*OV

2.4 Research questions

Children seem to show sensitivity to the properties of the language they are exposed to at the earliest observable stage of their syntactic productions (see VEPS by Wexler, 1998): infants raised in English-speaking environments produce the canonical VO order, while their Japanese peers produce the OV order. One question one may want to ask is whether this also holds in a language like Mandarin, in which there is no agreement, no overt case markers, and full topic-drop. Moreover, since frequency is argued to play a role in various approaches to language acquisition, in chapter 3 I consider some distributional facts about subjects and objects by focusing on the production of null/overt topics and the non-canonical *ba* construction by Mandarin-speaking infants and their caregivers and the effect of frequency in early language acquisition. In chapter 4 and chapter 5, I consider how early is the comprehension of canonical and non-canonical word orders in Mandarin, and whether infant comprehension is subject to crosslinguistic variation. The answers to these questions have implications for the usage-based approaches (Abbot-Smith & Tomasello, 2006; Tomasello, 2003) and the variational model (e.g. Yang, 2002, 2004) to language acquisition.

3 Spontaneous production of Mandarin word order

3.1 VEPS and the variational approaches to learnability

UG is thought to provide universal categories and features, which is genetically determined, part of the species endowment (Berwick, Chomsky, & Piattelli-Palmarini, 2013), and infants from different languages produce correctly the word order at the earliest observable stage (see evidence from Pierce, 1992 in early French and Poeppel & Wexler, 1993 in early German). In line with this, Wexler (1998) proposes the Very Early Parameter Setting (VEPS) hypothesis:

(44) Very Early Parameter Setting (Wexler, 1998):
Basic parameters of verb movement, e.g. V to I, V to I to C, are correctly set at the earliest observable stages (i.e. at the beginning of production of multiple word combinations, around 1;6).

VEPS is compatible with the triggering model (Gibson & Wexler, 1994) which claims that all children start out with either parameter setting, and, based on certain information in the input data, children can eventually change this initial parameter setting or consistently apply the initial value if it allows them to parse sentences. The algorithmic formulation is as follows:

(45) Given an initial set of values for n binary-valued parameters P, the learner,

a) Upon receiving an incoming sentence S, analyses S with P.

b) If successful, P remains unchanged; returns to a).

c) If failure, then

- changes the value associated with P, obtaining a new parameter setting P_n
- analyses S with P_n
- If successful, adopts P_n
- otherwise retains P; returns to a).

According to Gibson and Wexler (1994), the learner should have an initial hypothesis about the parameter setting; as stated by Wexler (2011) there is maybe some kind of biological mechanism which pushes "some parameters to be set first, refusing to let them change when other input comes in" (p. 94-95). Furthermore, the triggering model is error-driven, in that "the learner does not attempt to change her hypothesis as long as the current input sentence can be syntactically analysed" (Gibson & Wexler, 1994, p. 410). In addition, learning is online and conservative, in that a grammar G_n(s) is very close to the previous G, differing by only one parameter value.

As for the initial hypothesis and the ordering of parameter setting, Roberts (2016) and Roberts and Holmberg (2010) established a hierarchy. For example, based on whether there are ϕ-features in the relevant functional heads (in D and T), null-argument system follows a NO > ALL > SOME pattern (see table 2). Thus, they considered Chinese-like topic drop languages to be at the top of the hierarchy, which is default, followed by Spanish-like consistent null subjects, Finnish-like partial null subjects and English-like non-null subject languages.

Table 2: Scale of 'liberality' for null arguments (from (Roberts, 2016))

Rich D	Rich agreement	Types
-	-	Discourse NSLs
+	+	Consistent NSLs
-	+	Partial NSLs
+	-	Non-NSLs

Unlike deterministic models like the triggering model, the proposal of Yang (2002, 2004), on the other hand, is that learners do not select any parametric values as their initial guess, since both values are available. Yang (2002, 2004) regards language acquisition as the consequence of a population of competing grammars, whose distribution changes in response to the input presented to the learner. Hence, in this model (see (46)), the rise of the target grammar is gradual, which means that non-target grammars coexist for a while before they are eliminated. The procedure can be summarised as follows (from Yang, 2002, p. 26-27):

(46) For an input sentence S, the child:

i) with probability P_i selects a grammar G_i,

ii) analyses S with G_i,

iii)
- If successful, reward G_i by increasing P_i,
- otherwise punish G_i by decreasing P_i.

Therefore, in such model, the following predictions are made regarding acquisition. First, learners start with two parameter values, and the learning process is a trial-and-error search. Second, the behavior pattern takes the shape of a gradual curve, that is, abrupt changes in children's linguistic expressions should not be observed. The speed with which the target grammar rises to dominance is determined by the frequency of disambiguating evidence in the input.

Recently, Yang (2005, 2016) proposes a mathematical account for generalization during language acquisition named the Tolerance Principle, which seeks to define how many exceptions can a productive rule tolerate without the learner deciding to abandon it as unproductive. The central claim of that analysis is that to form a productive rule R, the number of exceptions (e) is less than the number of items (N) divided by the natural log of the number of items[13].

However, the Tolerance Principle may work well in morphology as to capture the English 'add-ed' rule for verb past tense (Yang, 2016), but it is unclear how it can apply to the acquisition of syntactic structure. One may consider the '*ba* + object' as some kind of chunk or morphological rule, perhaps analogous to 'verb + ed' in English. However, as discussed in section 2.3.3.1, the *ba* construction is a syntactic phenomenon, so if its acquisition follows from a morphological rule, then when does it become a syntactic rule? There is a problem of continuity here.

3.2 Aim of the present study

There is evidence that children show sensitivity to the properties of the language they are exposed to at the earliest observable stage of their syntactic productions. However, in the case of Mandarin, which allows NP arguments to be null, the input might be underspecified for

[13]The solution can be expressed as follows (cf. Yang, 2016):

(i) Tolerance Principle:
R is productive if and only if e $\leq \theta_N$ = N/In(N).

children to set correctly word order parameters. The aim of the present study is, first, to consider some distributional facts about null subjects and null objects and the canonical *ba* construction in Mandarin and its production in both children and adults, and, second, to evaluate Yang's variational model (2002; 2004), which highlights the importance of input frequency[14].

To do so, I conducted an analysis of the speech of Mandarin-speaking children from the corpus of Zhou (2001) and, since there were no subjects of age 6 and above in that corpus, five additional children from Chang's corpus (1998) were also included. Zhou's corpus includes children recruited in Nanjing, while Chang's children were recruited in Taiwan. They were all from Mandarin-speaking families as parents speak Mandarin Chinese to their children in everyday life. All transcripts can be found in the Chinese subset of the CHILDES database (MacWhinney, 2000). All children in this combined corpus were between the ages of 1;2 and 6;5 and were typically developing children. I selected my samples randomly. The corpus includes a total of 4624 child utterances and 8810 utterances from their adult interlocutors. Details on the child subjects can be found in Table 3, which includes information about the age and the MLU of each subject. And adult data sources are in Table 4.

[14]This chapter, coauthored with Anna Gavarró, was originally published in: Zhu, J.T., & Gavarró, A. (2019). Testing language acquisition models: Null and overt topics in Mandarin. *Journal of Child Language*, 46(4), 707-732. (DOI: https://doi.org/10.1017/S0305000919000114)

Table 3: Child subjects in the corpus

File	Subject	Age	MLU
cs14h.cha	Yangfan	1;2	1.357
cs14e.cha	Yijia	1;2	1.289
cs14b.cha	Liuxinyu	1;2	1.25
cs14g.cha	Xuyang	1;2	1.053
id14m.cha	Haohao	1;2	1.463
cs20i.cha	Xue'er	1;8	1.595
cs20c.cha	Wenwen	1;8	1.052
cs20g.cha	Jiangweiying	1;8	2.889
cs20d.cha	Qinlong	1;8	1.152
id20m.cha	Haohao	1;8	3.351
cs26i.cha	Majunhua	2;2.15	3.221
cs26b.cha	Liuzonghao	2;2.22	3.071
cs26j.cha	Shixintong	2;2.24	2.581
cs26c.cha	Shixuchen	2;2.26	1.466
id26m.cha	Haohao	2;2	3.558
cs32f.cha	Marui	2;8	3.514
cs32h.cha	Limanli	2;8	2.946
cs32c.cha	Chenzihui	2;8	2.574
cs32b.cha	Liyan	2;8	3.434
id32m.cha	Houhou	2;8	3.493
cs36fa07.cha	Wangyue	3;0	2.366
cs36fa08.cha	Zhouxinyuan	3;0	3.019
cs36fa09.cha	Hanjiaqi	3;0	2.293
cs36fb17.cha	Lishasha	3;0	2.12
cs36fb19.cha	Chenxixian	3;0	2.202
cs42mb12.cha	Zoushupeng	3;6	2.434
cs42ma03.cha	Zhangyuxuan	3;6	2.648
cs42fb20.cha	Caitianqi	3;6	2.317
cs42fb17.cha	Guohaohao	3;6	2.609
cs42fa10.cha	Chenziwei	3;6	2.747
cs48mb11.cha	Lijinghao	4;0	2.797
cs48fb16.cha	Chenxiaorong	4;0	3.213
cs48fa06.cha	Tangyi	4;0	2.377
cs48mb15.cha	Majunwei	4;0	2.722
id48m.cha	Haohao	4;0	4.278
cs60fa10.cha	Xuqucheng	5;0	2.325
cs60mb11.cha	Xuhao	5;0	3.736
cs60fb18.cha	Luohong	5;0	3.508
cs60fa09.cha	Yaoyifei	5;0	2.49
cs60fa06.cha	Dingyueying	5;0	3.46
cs66fa07.cha	Shikeyu	5;6	3.64
09.cha	Anxiang	5;7	4.761
10.cha	Xier	5;9	5.136
07.cha	Geli	5;11	7.092
cs72fa10.cha	Sunruiqi	6;0	3.573
05.cha	Lanxin	6;2	6.051
06.cha	Dezhi	6;5	7.810

Table 4: Adult subjects in the corpus

File	Identification
cs14h.cha	Mother
cs14e.cha	Mother
cs14b.cha	Mother
cs14g.cha	Mother
id14m.cha	Mother
cs20i.cha	Mother
cs20c.cha	Mother
cs20g.cha	Mother
cs20d.cha	Mother
id20m.cha	Mother
cs26i.cha	Mother
cs26b.cha	Mother
cs26j.cha	Mother
cs26c.cha	Mother
id26m.cha	Mother
cs32f.cha	Mother
cs32h.cha	Mother
cs32c.cha	Mother
cs32b.cha	Mother
id32m.cha	Mother
cs36fa07.cha	Mother
cs36fa08.cha	Mother
cs36fa09.cha	Mother
cs36fb17.cha	Mother
cs36fb19.cha	Mother
cs42mb12.cha	Mother
cs42ma03.cha	Mother
cs42fb20.cha	Mother
cs42fb17.cha	Mother
cs42fa10.cha	Mother
cs48mb11.cha	Mother
cs48fb16.cha	Mother
cs48fa06.cha	Mother
cs48mb15.cha	Mother
id48m.cha	Mother
cs60fa10.cha	Mother
cs60mb11.cha	Mother
cs60fb18.cha	Mother
cs60fa09.cha	Mother
cs60fa06.cha	Mother
cs66fa07.cha	Mother
09.cha	Observer/EXP
10.cha	Observer/EXP
07.cha	Observer/EXP
cs72fa10.cha	Observer/EXP
05.cha	Observer/EXP
06.cha	Observer/EXP

3.3 Method

For the purpose of this study, children were divided into nine age groups: the 1;2 age group with infants at only 1;2, the 1;8 age group with infants at 1;8, the 2;2 age group, the 2;8 age group, the 3;0 age group, the 3;6 age group, the 4;0 age group, the 5;0 age group and the 5;5–6;5 age group (with five speakers for each age group, except for the last group ranging from 5;5 to 6;5, which included seven speakers) as shown in table 5.

Table 5: Child subjects with age, mean MLU and Standard Deviation.

Age group	**Number of subjects**	**Mean MLU**	**Standard Deviation**
1;2	5	1.28	0.15
1;8	5	2.01	1.05
2;2	5	2.78	0.81
2;8	5	3.19	0.42
3;0	5	2.40	0.36
3;6	5	2.55	0.17
4;0	5	3.08	0.73
5;0	5	3.10	0.65
5;5-6;5	7	5.44	1.63

The analysis of the child and adult production was carried out manually unless otherwise indicated. With respect to subject production, I calculated the proportion of [± Null] subjects in topicalized structures; I found only two child-directed embedded sentences, with only one potential instance of a null *pro* subject – which was excluded from the recounts as involving another parameter (see section 2.3.1).

The proportion of [± Null] objects was calculated using a similar method. However, since all null objects are null topics, following Huang (1984a), I considered all [± Null] objects, including the *ba* construction.

In addition, when calculating null objects, following Huang's Phrase Structure Condition (1984b) and the criteria of Packard (2000), I excluded VO compound words as V+O structures. For example, *shui-jiao* in (47) is a compound word (literally consists of a verb *shui* 'sleep' and a noun *jiao* 'sleep'), while in (48) is a V+NP structure.

(47) Ta shui-jiao le.
he sleep PERF
'He has slept.'

(48) a. Ta shui le liang ge xiao-shi de jiao.
he sleep PERF two CLASS hour DE sleep
'He slept for two hours.'

b. Jiao, wo xiang ta shi hui shui de.
sleep, I think he be will sleep DE
'Sleep, I think he will take.'

Shui-jiao 'sleep' in (47) is a word because it may not be followed by an object. In contrast, *shui-jiao* 'sleep' in (48a) and (48a) is a transitive phrase, because the relation between the verb *shui* and object *jiao* is clearly syntactic as the object is modified by the expression of time or has been topicalized, none of which are lexical operations.

Furthermore, the transitive sentences were picked based not only on the main verbs used, but also the context in which the sentences in question were actually produced, because a verb can be optionally transitive. Thus, grammaticality and acceptability were judged with respect to the given discourse context. I excluded any utterance that appeared to be an immediate imitation from earlier in the same transcript (Demuth, 1996) or "chunks" that frequently occur in the input (Tomasello, 2003). Such utterances were excluded on the grounds that they do not reliably indicate the child's own grammatical knowledge.

Since efficient automated analysis tools are now readily available, I used CLAN tools (MacWhinney, 2015), in particular the command kwal +t*CHI +s“ 把 ” @ and kwal +t*MOT +s“ 把 ” @ to cull all of the utterances that contained the *ba* construction in the children’s and adults’ spontaneous speech.

3.3.1 Data treatment and analyses

First, descriptive statistics are reported. Then, a Poisson Regression model for the null subject and null object is set up. An offset term and age group as covariate were included. The offset term in the Poisson regression model is used to model rates rather than counts; this is necessary for instance when individuals are not followed for the same amount of time. In this paper the offset term for the null subject model was total subjects = null subjects + overt subjects and total objects = null objects + overt objects for the null object model. Results from the model are reported below with the estimated proportion and its confidence interval. All results were obtained using R software (version 3.5.2). R packages used in the analysis were stats, emmeans and RcmdrMisc (R Development Core Team, 2015).

3.4 Results

First I report the overall percentage of subject omission in the corpus; I turn to the omission of canonical objects and in the *ba* construction in the last two subsections.

3.4.1 Results for subject production and omission

Out of all sentences, the mean percentage of sentences with null subjects produced by the Chinese children was 48.98 (Standard Deviation, *SD* = 24.86; Standard Error mean, *SE* = 3.62), while it was 49.83 (*SD* = 10.82; *SE* = 1.58) for the Chinese adults. I analyzed whether there were differences in the proportion of subject omission as a function of age and found that initially (in the period starting at age 1;2) all five children studied started out by omitting all subjects (M = 100, *SE* = 0). At 1;8, the production of subjects by children increased, and in addition children showed an early sensitivity to language-specific properties of the input data with a subject being produced when required, as in (49), and a null subject produced when the context favored it (50).

(49) Wo yao he cha.
I want drink tea
'I want to drink tea.' (Jiang Weiying, 1;8)

(50) a. MOT: Yao bi ma?
want pen Q
' Do you want a pen?'

b. CHI: Yao.
want
'(I) want (it).'
(Jiang Weiying, 1;8)

The number and percentage of null subjects found in child- and child-directed speech (CDS) appear in Table 6.

The developmental curve of subject omission for the nine age groups is shown in Figure 1.

Table 6: Percentage and number of null subjects over all sentences (only sequences with a verb included).

Age Group	Null subject (Child) %	N	Null subject (CDS) %	N
1;2	100	18/18	55.66	177/318
1;8	58.43	52/89	55.52	307/553
2;2	51.05	97/190	37.31	175/469
2;8	61.67	74/120	56.67	187/330
3;0	46.15	54/117	52.67	256/486
3;6	42.22	57/135	53.05	270/509
4;0	58.90	129/219	44.23	230/520
5;0	35.10	119/339	47.41	238/502
5;5-6;5	49.49	97/196	47.27	52/110
Mean	**48.98**	697/1423	**49.83**	1892/3797

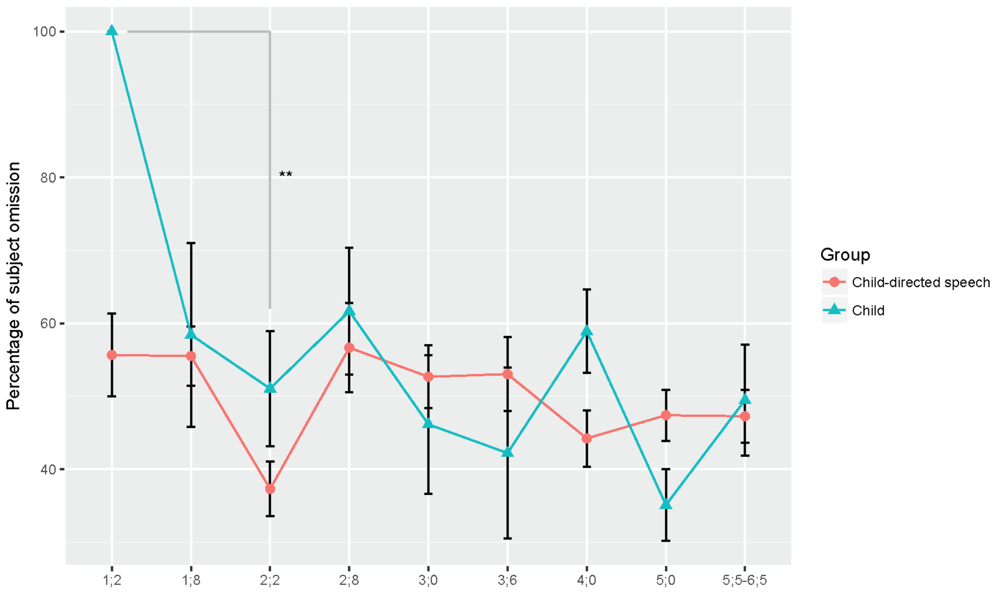


Figure 1: Developmental curve of subject omission in child- and child-directed speech.

The results indicate that the omission of subjects by children stays around 50% from 1;8 to 6;5 (the oldest age considered). Moreover, a Poisson regression model with an offset term was established to predict the number of null subjects produced by children based on age. In this model, I excluded the 1;2 age group from the developmental curve of null subjects because children in this age group had not yet reached the two-word stage (their MLU was below 1.75), so it is virtually impossible for them to produce the arguments of an overt verb, or at least to do so in any systematic way. Thus I report the estimated proportion of null subjects and its 95% confidence intervals (CIs) for children from 1;8 to 6;5 and from adult as shown in table 7.

Table 7: Estimated mean of null subjects and its 95% confidence intervals.

Age group	**Mean(%)**	**Lower Confidence Limit**	**Upper Confidence Limit**
1;8	58.43	44.52	76.67
2;2	51.05	41.84	62.29
2;8	61.67	49.10	77.45
3;0	46.15	35.35	60.26
3;6	42.22	32.57	54.74
4;0	58.90	49.57	70.00
5;0	35.10	29.33	42.01
5;5-6;5	49.49	40.56	60.39
Adult	49.83	47.63	52.13

To examine if there is a significant difference between ages 1;8 (M = 58.43%, $SE = 12.60$) and 2;2 (M = 51.05%, $SE = .79$) with respect to the percentage of null subjects produced by children, I ran a post-hoc test. Though the descriptive statistics showed a decline, these differences were not statistically significant ($t(8) = 1.84$, $p = .104$), which means there is no abrupt change in children's performance.

It is possible that the children omitted subjects because they imitate null subjects in the input. I calculated the percentage of null subjects in all sentences heard by the children (including fragments, imperatives, etc. N = 1892), which constitutes 21.5% of the total number of sentences in the corpus (N = 8810). I also examined the incidence of null subjects in adult use at each point in the child's development represented in Table 6 and Figure 1 above: the post-hoc test shows that there is no significant difference between children and adult performance by the period starting at age 1;8 (t(4.83) = – 2.04, p = .099). Moreover, a Poisson regression model shows that from 1;8 (the youngest stage considered), age has no statistically significant effect on the production of null subjects by the children. Moreover, from 1;8, children's performance on null subjects is not significantly different from the adult's production, which confirms again my conclusion that the children's production of null subjects has become adult-like by 1;8, that is, quantitatively speaking, by age 1;8 the child production of null subjects has become adult-like. Although the children's subject-drop rate does not approach that of the adults before age 1;8, at 1;8 the parameter value that allows null subjects seems to be set. Besides, there is no systematic decrease in the rate of null subjects in the adults' speech that parallels the decrease in the children's use of null subjects, as might be expected if children merely imitated the input.

3.4.1.1 Results across MLU in subject omission

In order to determine whether there is any relationship between the null subject phenomenon and the child's linguistic development, the percentage of null subjects was recalculated on the basis of the children's MLU. Based on Brown's stages of language development (Brown, 1973), I divided child production into six MLU groups, as in Table 8. A strong positive correlation was found between MLU and chronological age of my Chinese children (r = .745, $p < .001$).

Table 8: Percentage of null subject sentences on the basis of MLU, standard deviation (SD) and standard error mean (*SE*) at each stage.

MLU Stage	**MLU range**	**Null subject**	***SD***	***SE***
I	1.0–1.75	93.18%	11.11	3.70
II	1.75–2.25	60.42%	21.07	14.90
III	2.25–2.75	49.27%	23.92	6.63
IV	2.75–3.5	49.21%	17.33	5.23
V	3.5–4.0	36.57%	20.57	8.40
VI	4.0+	73.56%	8.09	3.30

I excluded those children whose MLU was below 1.75 (Brown's MLU stage I) from the curve (see Fig. 2), since they had not fully reached the two-word stage; because subject production requires the presence of a verb as well, subject omission before MLU 2 may simply result from processing factors – I take those to be responsible for a low MLU in early production. The curve confirms what I found when considering age. There is no statistically significant difference from MLU 1.75–2.25 (Brown's Stage II) to 2.25–2.75 (Stage III) and there is no significant difference between children and adult performance at MLU Stage II regarding subject drop, which means that before MLU 2.25 their use of null subjects has already become adult-like.

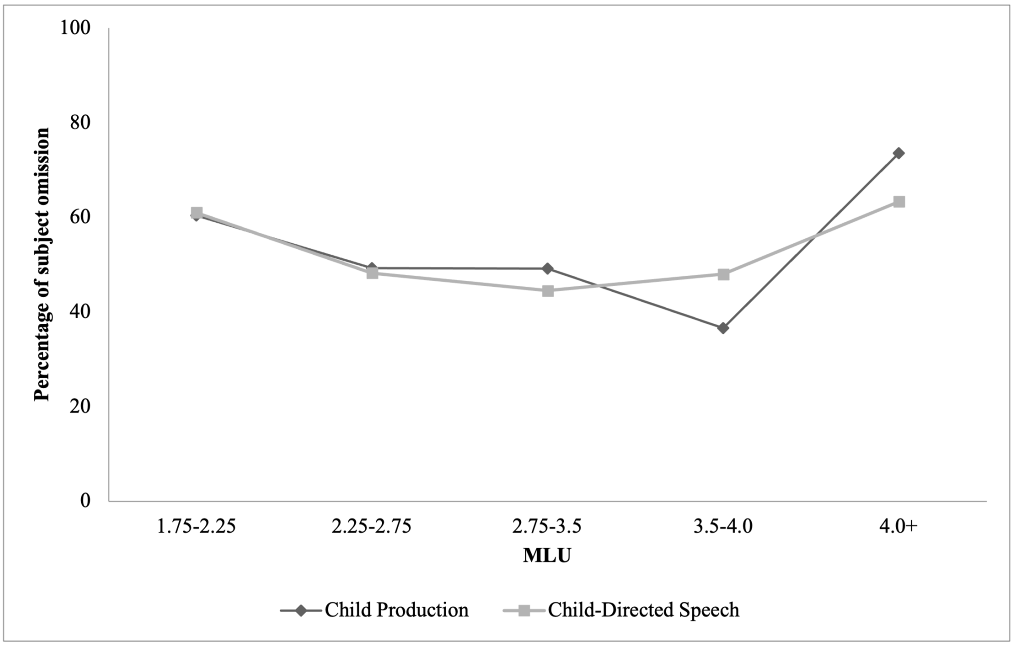


Figure 2: Developmental curve of null subjects by Chinese children by MLU and in child-directed speech.

In sum, whether divided by age or by MLU, children set the correct parameter value at a very early age (at age 1.8 or before MLU 2.25). This is also consistent with individual data, that is, no variability among the children was found (examples from five individual productions can be found in Appendix A.1). All children seemed to have correctly set the parameter value for null subjects by age 1;8.

3.4.2 Results for object production and omission

The mean percentage of sentences with null objects produced by the Chinese children was 33.19 ($SD = 26.67$; $SE = 3.96$), while it was 34.42 ($SD = 16.13$; $SE = 2.24$) for the Chinese adults.

The results for child- and child-directed speech appear in Table 9 and

Figure 3. For statistical validity, as in the case of subjects, I also excluded children in the age group of 1;2 from my calculation, since at this age children omitted objects 100% of the time.

Table 9: Percentage and number of null subjects found (only transitive sentences considered).

Age Group	Null object (Child) %	Null object (Child) N	Null object (CDS) %	Null object (CDS) N
1;2	100	14/14	43.05	65/151
1;8	42.86	15/35	38.67	87/225
2;2	34.57	28/81	26.73	50/187
2;8	41.77	33/79	32.93	54/164
3;0	30.59	26/85	31.87	109/342
3;6	31.31	31/99	28.21	110/390
4;0	28.66	45/157	40.36	134/332
5;0	31.67	57/180	37.35	121/324
5;5-6;5	29.80	59/198	32.26	30/93
Mean	**33.19**	308/928	**34.42**	760/2208

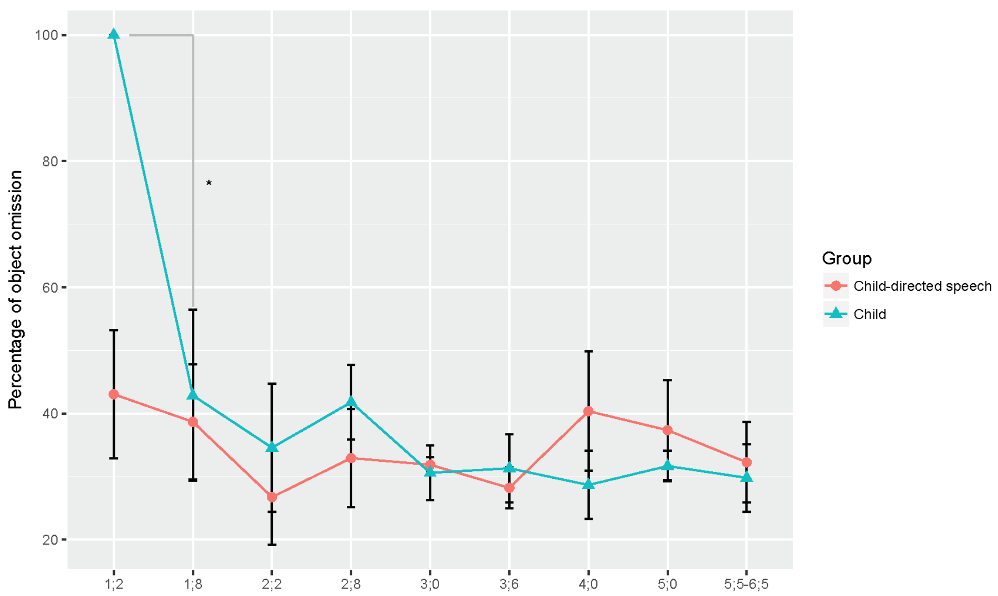


Figure 3: Developmental curve of object omission in child- and child-directed speech.

The results in Figure 3 indicate that null object production by children stays around 35% from 1;8 to 6;5. I also report the estimation of null objects and its 95% CIs in Table 10 below. Again, a post-hoc test was run to verify if there is a significant difference between ages 1;8 (M = 42.86%, *SE* = 13.56) and 2;2 (M = 41.76%, *SE* = 10.15), but none was found ($t(8) = .58$, $p = .578$).

Table 10: Estimated mean of null objects and its 95% confidence intervals.

Age group	**Mean(%)**	**Lower Confidence Limit**	**Upper Confidence Limit**
1;8	42.86	25.84	71.09
2;2	34.57	23.87	50.06
2;8	41.77	29.70	58.76
3;0	30.59	20.83	44.93
3;6	31.31	22.02	44.53
4;0	28.66	21.40	38.39
5;0	31.67	24.43	41.05
5;5-6;5	29.80	23.09	38.46
Adult	34.42	32.06	36.96

The percentage of null objects constitutes 8.6% over the total number of sentences in the corpus (including fragments). Following the same methodology used with null subjects, I also examined the incidence of null objects in adult use at each point in child development: first, again, there is no systematic decrease in the rate of null objects in the adults' speech that parallels the decrease in the children's use of null objects, as might be expected if children merely imitated the input. Second, there is no significant difference between child and adult performance by the period starting at age 1;8 ($t(8) = -1.1$, $p = .304$) and the Poisson model also confirms my finding; that is, quantitatively, as early as 1;8 the child production of null objects cannot be distinguished from that of Chinese adults.

In the qualitative study, I counted all verbal utterances in the period starting at 1;8 for which the object is obligatory based on adults' judgements in the specific discourse context. The use of overt objects is target-consistent. This is illustrated in (51) and (52).

(51) a. MOT: Gao-su a-yi xiao-bai-tu chi shen-me?
tell auntie little-rabbit eat what
'Tell the auntie what rabbits eat.'

b. CHI: Chi luo-bu.
eat radish(*Ø)
'(They) eat radish.'

(Xu'er, 1;8)

(52) a. MOT: Ni yao zuo ha?
you want do what
'What do you want to do?'

b. CHI: Wo yao na zhe-ge.
I want take this
'I want to take this.'

(Jiang Weiying, 1;8)

The use of null objects in that period is also target-consistent. All children produced null objects just like adults, as exemplified in (53) and (54).

(53) %act: MOT is teaching how to pull the pen cup out.

Nai-nai ba.
grandma pull-out
'Grandma pull (it) out.'

(Xu'er, 1;8)

(54) %act: CHI is taking the pen out.

Ma-ma	na	hao.
mom	take	good

'Mom, take (it) well.'

(Jiang Weiying, 1;8)

Utterances (53) and (54) are felicitous with the reference of the null objects determined by the discourse topic. In this regard, the children's null object use at age 1;8 has already converged to the adult grammar (more individual data can be found in Appendix A.2).

3.4.2.1 Results across MLU for object production and omission

The results in Table 11 show that all the children start out by omitting objects very frequently. This result is hardly surprising, given that children at Stage I have a very low MLU, and an overt object (just like an overt subject) requires a two-word utterance at least.

Table 11: Percentage of sentences with a null object based on MLU.

MLU Stage	**MLU range**	**Null object**	***SD***	***SE***
I	1.0–1.75	85.19%	22.74	7.58
II	1.75–2.25	34.33%	9.68	5.39
III	2.25–2.75	31.21%	10.43	2.82
IV	2.75–3.5	33.42%	12.60	3.52
V	3.5–4.0	26.93%	13.52	8.56
VI	4.0+	37.56%	34.43	14.06

If I exclude the first stage from the calculation, the curve by MLU in Figure 4 confirms what I found when considering age. There was no statistically significant difference between MLU 1.75–2.25 (Stage II) to MLU 2.25–2.75 (stage III) in children's production with respect to object drop and, from MLU 1.75–2.25, the children's performance regarding null objects has become target-like.

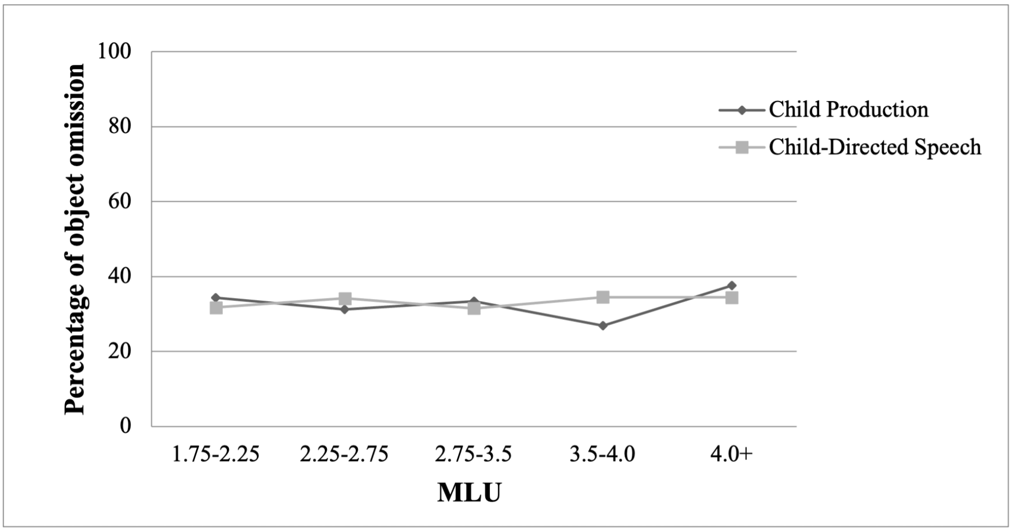


Figure 4: Developmental curve of null objects by Chinese children by MLU and in child-directed speech.

In sum, whether divided by age or by MLU, children's production of null objects is adult-like at a very early age (age 1.8 or before MLU 2.25). It seems unlikely that children omit objects because they mimic the input. If this were the case, I should have found a systematic change in the input (adult child-directed speech) parallel to the change in the output (child speech). However, this was not the case. Thus, different performance among the nine age groups cannot be attributed to the different rate of object drop in the input available to the child.

3.4.2.2 Results for additional ages between 1;2 and 1;8

Since there is a relatively abrupt change between 1;2 and 1;8, it is possible that more evidence of gradual change might have been found if additional ages between 1;2 and 1;8 were examined. Thus, I conducted a follow-up study using the corpora of Xu/Min/Chen (seven transcripts from two participants, Chen, 2008), Tong (one transcript from one participant, Deng & Yip, 2018) and TCCM (seven transcripts from four participants, Cheung et al., 2011) from CHILDES. Details of the data are reported in table 12 (data belonging to each individual corpus can be found in CHILDES, TalkBank/CHILDES [Index to Corpora: Chinese]).

Table 12: Data for additional ages.

Corpus	**Number of children**	**Age range**	**Region**
Xu/Min/Chen	2	0;11-1;7	Beijing
Tong	1	1;7	Shenzhen
TCCM	4	1;5-1;7	Taiwan

The following example (55) shows that null object production is set correctly even at age 1;3.

(55) a. FAT: Mei jiao guo ta ma?
no call EXP he Q
'Did (you) call him?'

b. CHI: Mei jiao guo (Ø) .
no call EXP
'(I) didn't call (him).'

(Ma, 1;3.15)

(56) a. MOT: Ni yao shen-me?
you want what
'What do you want?'

b. CHI: Yao bu-ding.
want pudding
'(I) want pudding.'

(Yang, 1;5)

The examples above support a parameter-setting model over an input-driven learning model, with a null topic being produced when the context allows for it (55), and an overt object produced when required (56).

3.4.3 Results for object production and omission in the *ba* construction

Now I turn to object production in the *ba* construction. I used CLAN tools (MacWhinney, 2015) to cull all of the utterances that contained the *ba* construction in the children's and adults' spontaneous speech (see section 3.3 for the code).

Both in young children and adult speech its frequency is relatively low according to my data. In young child production, the total number of sentences involving the *ba* construction in my corpus is 63 and in adult speech it is 298; the overall percentage of the *ba* construction against transitive sentences is 6.79% ($SE = 1.14$) in child production, while in adult speech it is 13.5% ($SE = 1.42$). I also calculated the frequency of adult usage of the *ba* construction in the input (including fragments), which constitutes 3.1% of all utterances. Table 13 reports the results in detail.

As shown in the developmental curve in Figure 5, the *ba* construction is first attested at 1;8 and reaches quantitative adult levels of production by 4;0.

Table 13: Percentage and number of *ba* constructions against transitive sentences.

Age Group	*ba* construction (Child)		*ba* construction (CDS)	
	%	N	%	N
1;2	0.00	0/14	17.88	27/151
1;8	2.86	1/35	22.22	50/225
2;2	2.47	2/81	22.99	43/187
2;8	2.53	2/79	9.76	16/164
3;0	3.53	3/85	12.87	44/342
3;6	2.02	2/99	12.56	49/390
4;0	10.19	16/157	11.14	37/332
5;0	9.44	17/180	6.79	22/324
5;5-6;5	10.10	20/198	10.75	10/93
Mean	**6.79**	63/928	**13.50**	298/2208

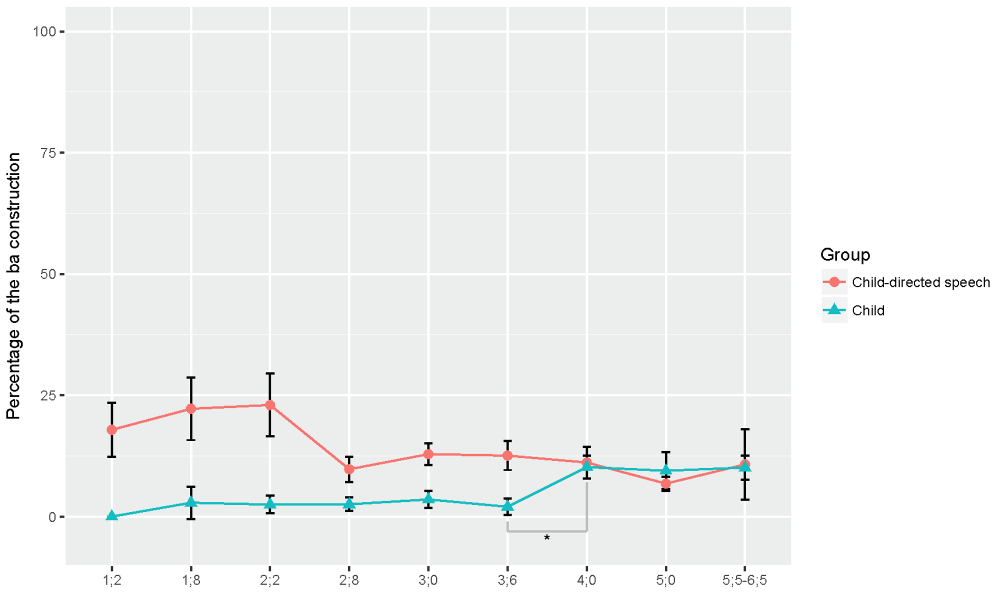


Figure 5: Developmental curve of the *ba* construction in child- and child-directed speech.

There was a significant change between 3;6 and 4;0 (t(8) = –2.56, p = .034) and it represented a large-sized effect (d = 1.62). However, the input received did not present a significant change (t(8) = 1.07, p =

.315), suggesting that the change could not be the result of mimicking the input.

The percentage of the *ba* construction against transitive sentences based on the children's MLU was also calculated as seen in Table 14; the result is also shown in the developmental curve in Figure 6 (again I excluded the first MLU stage):

Table 14: Percentage of the *ba* construction against transitive sentences on the basis of MLU.

MLU Stage	MLU range	*ba* construction	*SD*	*SE*
I	1.0–1.75	0.00%	0.00	0.00
II	1.75–2.25	0.00%	0.00	0.00
III	2.25–2.75	3.00%	4.64	1.29
IV	2.75–3.5	6.03%	6.96	1.71
V	3.5–4.0	10.91%	6.68	6.83
VI	4.0+	14.22%	13.98	5.71

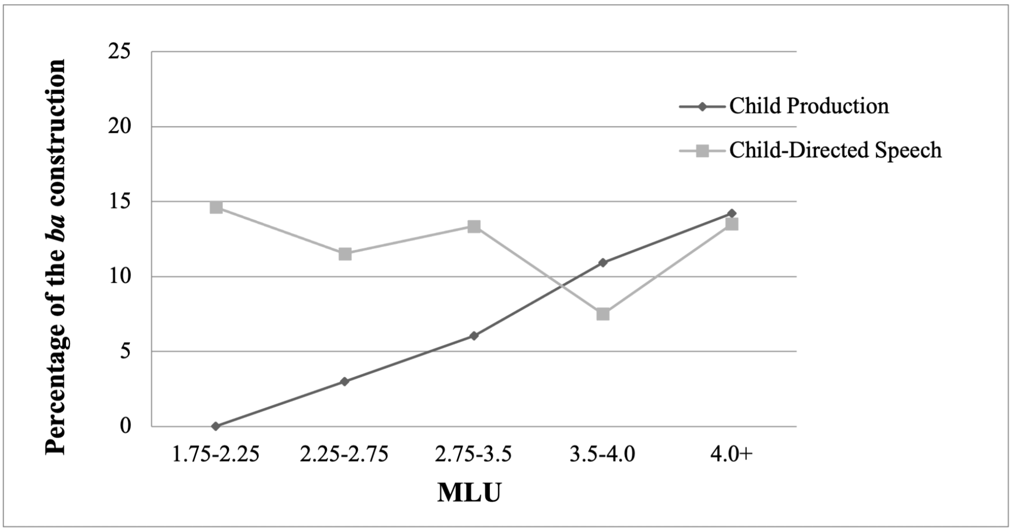


Figure 6: Developmental curve of the *ba* construction by Chinese children by MLU and in child-directed speech.

As can be observed, although the Mandarin-speaking children did not produce the *ba* construction during the first two stages studied (MLU

1.0–2.25), after Stage II, as their MLU increased, the mean percentage of sentences with the *ba* construction increased. By MLU 3.5 their use of the *ba* construction is approaching that of Chinese adults (t(10) = .15, p = .883). Most importantly, all the children whose productions I analyzed produced an overt object (i.e. post-*ba* NP) in the *ba* construction from the outset (63 sentences with *ba* + NP and no sentences with *ba* followed by object omission). Examples are provided in (57).

(57) a. Ba da-hui-lang gan zou le.
BA wolf drive away PERF
'(I) drove the wolf away.' (Haohao, 1;8)

b. Bang wo ba ta na chu-lai.
help I BA it take out
'Help me to take it out.' (Liuzonghao, 2;2)

c. Ta ba ping-guo reng diao le.
he BA apple throw away PERF
'He trew away the apple.' (Marui, 2;8)

d. Ni ba zhe-ge na zou, ba na-ge na chu-lai.
you BA this take away BA that take out
'You take this away, and take that out.' (Lianlian, 3;0)

In the *ba* construction, an overt object is expressed in 100% of utterances by infants; the observation that children's *ba* productions reliably included an overt object is also supported by a paired *t*-test, which indicates that the frequency of production of an overt object in the presence of *ba* is significantly greater than its frequency when *ba* is absent (t(41) = -10.42, $p < .001$, $d = 3.25$). Although in the corpus I did not find any use of topic markers after the post-*ba* NP by children or by adults, the contexts in which children used post-*ba* NPs are compatible with a contrastive topic interpretation (see 57d). In addition, the fact that all children produced post-*ba* NPs just like adults shows that, if I adopt the standard criterion (e.g. Brown, 1973; Thornton & Tesan, 2007) that 90% adult-like usage in obligatory contexts indicates that a grammar structure has been acquired, the Chinese *ba* construction is an early acquisition (attained by 1;8), although the frequency of occurrence of this structure is low in my corpus (only 3.1%).

Most importantly, the group results are corroborated by individual data. Over the course of language acquisition, the individual use of null objects became adult-like by age 1;8 (see Appendix A.2 for more examples). Regarding the *ba* construction, target-inconsistent use was unattested for the entire data set.

3.5 Discussion

The results show that Mandarin-speaking children have a grammar that allows for null topics at an early age judging by their naturalistic speech, that is, roughly at the age of 1;8, or before MLU 2.25; the production of both null subjects and null objects by Chinese children is indicative of a target-like grammar. My study provides evidence that children acquiring Chinese are sensitive to topicalization at around 20 months of age. To my knowledge, there is no study of the early acquisition of topic markers in Mandarin, but the evidence presented here is consistent with the finding that Korean children are learning to understand and produce topic markers, including contrastive topic markers, before age 2;0 (Lee & Cho, 2009).

Let us turn to two predictions of Yang (2002, 2004)'s variational model. The input frequency of null objects (8.6%) is very low compared to that of null subjects (21.5%), which might imply a longer variational stage with null objects than with null subjects. French speaking children learn that French is a verb raising language by 1;8 according to Pierce (1989). Yang (2002), based on the CHILDES database, estimated that signature sentences for verb raising, which is expressed by the form V_{FIN} Neg/Adv, constitute 7% of all the sentences that children acquiring French hear. Thus, he concluded that the frequency of the signature for a parameter to be set early must be at least 7% of the input data. Yang (2002) also asserted that the amount of evidence for null object utterances in the input is sufficiently high for children to acquire the Null Topic parameter very early based on the data in Wang, Lillo-Martin, Best, & Levitt (1992), and at first sight my data seem to support the quantitative analysis of Yang. Null subject sentences constitute 21.5% of the input data, and null object sentences constitute 8.2% of the input data, both percentages above the 7% threshold, thus guaranteeing early acquisition.

However, the *ba* construction, which requires an overt object, which in turn is constrained by another parameter, comprises just 3.1% of the input. According to Yang's calculation, this structure should be acquired late, because its frequency is lower than the threshold for an early acquisition (i.e. 7%). Yet, contrary to this prediction, children are able to correctly set the post-*ba* NP from the outset. The relatively high frequency of null objects in Chinese (34.42%) and the low frequency of the *ba* construction (3.1%) in the input should lead children to drop objects (i.e. post-*ba* NP) in this construction at least at the early stage. However, this is in disagreement with the fact that all the children in my study were able to produce the post-*ba* NP in a target-like fashion from the outset at age 1;8 (see (57a)), even if an NP in Chinese can be optionally dropped, which reveals that the *ba* construction is an early acquisition. There is no variational stage such as the one expected in the variational model.

Even if there is no evidence for a variational stage in child production, it is possible that input factors are responsible for the different rates and patterns of production as proposed by Yang (2002, 2004). Under the variational approach of probabilistic learning (Legate & Yang, 2007; Yang, 2016), if a grammar has a higher number of ambiguous forms, the child will take much longer to acquire that grammar. Thus, by calculating the proportion of [±Null Topic] subjects and objects in the input, we can predict the time to acquire a Mandarin-like [+Null Topic] grammar according to the variational model.

Based on that, I compared the overall rate of [±Null Topic] subjects and objects in adult production. The results of my re-counts are summarized in Table 15.

Table 15: Numerical advantage of [–Null Topic] grammar in adult production.

Rewards [+Null Topic] subject	1892/3797 (49.83%)
Rewards [–Null Topic] subject	1905/3797 (50.17%)
Rewards [+Null Topic] object	760/2208 (34.42%)
Rewards [–Null Topic] object	1448/2208 (65.58%)
Rewards ([–Null Topic] subject + [–Null Topic] object])	
– ([+Null Topic] object + [+Null Topic] subject)%	55.8% – 44.2% = 11.6%

I found 2652/6005 (44.2%) recorded sentences that unambiguously implicate a [+Null Topic] grammar. These are countered by 3353/6005 (55.8%) sentences recorded that, by virtue of being consistent with the [–Null Topic] grammar, may impede the acquisition of the target Chinese [+Null Topic] grammar. This means that the numerical disadvantage of the target null Topic grammar is 11.6%.

In the case of the *ba* construction, since an overt object is always required, there is no [+Null Topic] option and the child only has evidence to consider her grammar as non-null object. If I exclude this structure from the setting of the [+Null Topic] grammar due to its non-optionality (my calculation is conservative; in fact, one would argue that *ba* constructions are part of the input, and therefore part of the calculation), the numerical advantage of the target [+Null Topic] grammar is still low (i.e. lower than 7%).

One can find more or less the same numerical percentage in Legate and Yang (2007)'s work in the context of the acquisition of English [+Tense] grammar – see Table 16 below. The numerical advantage of the [+Null Topic] grammar is much lower than the advantage of the [+Tense] grammar of English (5.8%). Based on that, I expect, if we adopt Yang's model, that Chinese children will converge to a target parameter setting (that is, a Null Topic grammar) even much later than English-speaking children acquire a [+Tense] grammar (i.e. more than 3;5). However, both quantitative and qualitative data show that Chinese children are performing at ceiling before 3;5, which reveals that the implications of Yang's model are not fulfilled.

Table 16: Quantitative evidence in favor of the [+Tense] grammar and the reported duration of the RI stage in three languages (from Legate & Yang, 2007, p. 336).

Language	% [+T] – % [–T]	Duration
Spanish	60.2%	$\sim$ 2;0
French	39.6%	$\sim$ 2;8
English	5.8%	$>$ 3;5

Moreover, my findings show that the *ba* construction is found at a rate of 3.1% of all the sentences in the input; therefore, as discussed above, following the variational model, this structure should be acquired late. However, this is in disagreement with the fact that all the children in my study were able to produce the post-*ba* NP in a target-like fashion from the outset at age 1;8, which reveals that the *ba* construction is an early acquisition.

Admittedly, the conclusions I draw on the basis of naturalistic data depend on the number of data points considered; the corpus examined here is of considerable size, and the results are very consistent across ages. My sample size of 6,005 input sentences with [±Topic] is bigger than Legate and Yang (2007)'s sample of 2,226 sentences for Spanish and 2,231 for French, and smaller than their 70,047 sample for English. The evidence I have presented shows that the setting of certain parameters can be very quick, despite how much (or little) unambiguous evidence there is in the input. The setting of the *ba* construction also suggests that children are not simply setting parameters based on what appears most frequently in their input, which means that parameter setting cannot be explained simply by input frequency, at least not in a superficial sense, as observed by many others (Friedmann & Costa, 2011; Gavarró et al., 2010; Guasti et al., 2008; Roeper, 2007; Thornton & Tesan, 2007). At least, if frequency matters, it is not at 17 months or after.

On the other hand, if we assume with, for example, the triggering model (see Gibson & Wexler, 1994) that the child starts with a [-Null Topic] setting, upon encountering null topics in Chinese, the child will set the parameter correctly to the + value. The setting of the *ba* parameter (depending on the acquisition of a functional element) follows the same procedure. This could capture the facts I showed in the study. Alternatively, the acquisition of null topics could be explained by the parameter hierarchy of Roberts (2016) and Roberts and Holmberg (2010) following a NO > ALL > SOME pattern (repeated here as (58)), which may lead to a learning path as assumed by Biberauer and Roberts (2016).

(58) Chinese-like discourse NSLs > Spanish-like consistent NSLs > Finnish-like partial NSLs > English-like non-NSLs.

As we move successively down in the hierarchy, systems become more marked, and more computationally complex to learn; children only adopt such settings when forced to by Primary Linguistic Data (PLD) incompatible with the less marked setting. Thus, Mandarin-speaking children set the [+Null Topic] parameter[15] very quickly because it is at the top of the hierarchy, which is the option preferred by the learner and is assumed to represent their initial hypothesis.

[15] Here, the null topic is the same as discourse NSLs.

4 Early comprehension of Mandarin canonical word order

In this chapter, I report a study on the comprehension of canonical word order in Mandarin-speaking infants aged from 1;1 to 1;9, where they are still in the pre-productive stage of language acquisition. First, I introduce the methodology used in the experiments.

4.1 Intermodal Preferential Looking Paradigm using eye-tracking

The Intermodal Preferential Looking Paradigm (IPLP) is used for research on language development before children can talk. In the IPLP, infants' language comprehension is measured by their differential visual fixation accompanied by an auditory stimulus matching only one of the stimuli (Golinkoff et al., 1987; Hirsh-Pasek & Golinkoff, 1996). Golinkoff et al. (1987) was the first team using the IPLP to examine whether 16-month-olds could comprehend nouns (e.g., shoe vs. boat) and verbs (e.g., drink vs. blow). The results showed that infants at 16 months looked significantly longer at the objects or actions that matched the language they heard rather than at the scene that did not match. Then Hirsh-Pasek and Golinkoff (1996) used the same methodology to investigate whether 17-month-old infants use word order to find "who-did-what-to-whom".

In their paradigm, normally the infant is seated on a parent's lap in front of two televisions or one with a split-screen and a hidden camera below the television records infants' visual fixation. The disadvantage of this technique is that it often requires long hours of manual coding, and there is room for error in coding (Dalrymple, Manner, Harmelink, Teska, & Elison, 2018; Hessels, Cornelissen, Kemner, & Hooge, 2015).

In the last decades, eye-tracking has become a popular research tool in developmental cognitive neuroscience (see e.g., Aslin & McMurray, 2004; Aslin, 2012). Using an eye-tracker, rather than a camera, gives as output a gaze signal on a screen, which is obtained by directly filming the eyes of the participant through several (near-)infrared illuminators. Nowadays, eye trackers can automatically report gaze location up to a thousand times per second with no manual labor involved at all (Holmqvist et al., 2011) and provide more reliable and detailed results (Dalrymple et al., 2018; Hessels, Cornelissen, et al., 2015). Therefore, experimental techniques such as eye-tracking improve on previous methods as they allow for more natural processing and yield very rich temporal data. Compared to the traditional manual gaze coding method, eye-tracking is quite objective, efficient, and has relatively high temporal and spatial resolution (Dalrymple et al., 2018; Hessels, Cornelissen, et al., 2015).

In this framework, Franck et al. (2013) tested French-speaking infants of 19 months and their aim was to determine if infants were aware of the VO/OV contrast. In their experiment, they resorted to the WWO paradigm (Akhtar, 1999, see 1.3.1 for more details) in which children heard grammatical NP-V-NP and weird/ungrammatical NP-NP-V sequences with pseudo-verbs as in (59).

(59) a. Le lion poune le cheval.
DET lion PSEUDO-V DET horse
'The lion V-ed the horse.'

b. *Le vache le lion dase.
DET cow DET lion PSEUDO-V
'The cow the lion V-ed.'

Also, the distractor video critically differed from the study of Gertner et al. (2006): rather than illustrating reversed theta-roles, it illustrated a transitive action and the same action performed reflexively. The prediction is that if the adult-like abstract syntactic knowledge is available for infants from early on, they can distinguish between VO and OV sequences and realise that French is VO. By contrast, if word order is encoded as a lexical property, infants will show random behaviour in both conditions since experimental sentences contain pseudo-verbs.

The results indicated that infants only looked at the transitive scene when they heard the grammatical sequence, while they showed random behavior when they heard the ungrammatical sentences. The preference for the SVO interpretation of NP-V-NP sentences provides strong evidence that infants know that the NP following V is its object, while the preference of English infants reported by Gertner et al. (2006) could be due to a preference for SVO over OVS, in line with the quasi-universal SO order found across languages. Moreover, the lack of a preference for NP-NP-V sentences shows that the NP preceding V is not erroneously interpreted as its object. Similar results were obtained by Gavarró, Leela, Rizzi, & Franck (2015) in the same framework (see also Leela, 2016). They tested 20 infants aged 19 months exposed to an OV language with case-marking, Hindi-Urdu (see (60)), and the results show that infants can parse the grammatical SOV sequences as indicated by them looking significantly longer at the transitive video, but they failed to assign a consistent interpretation to the ungrammatical VSO order. Taken together, these experiments indicate that the parameter responsible for the VO/OV alternation is set correctly by 19 months regardless of lexical knowledge of the verb.

(60) a. Sher-ne ghode-ko khalaa-yaa.
lion-ERG horse-ACC PSEUDO-V-PERF
'The lion V-ed the horse.'

b. *Choona gaay-ne sher-ko.
PSEUDO-V-PERF cow-ERG lion-ACC
'The cow the lion V-ed.'

However, a number of studies of children's early productions conducted in Mandarin have argued to provide counter-evidence to that conclusion. Chan, Lieven, & Tomasello (2009) used an act-out task and found that Cantonese children did not choose the first noun as AGENT in the canonical SVO sentences containing pseudo-verbs at above-chance levels until 3;6. Examples of test sentences are provided in (61).

(61) a. Zyul-zai tam joeng-zai.
pig PSEUDO-V sheep
'The pig V the sheep.'

b. Maa-zai tam din-waa.
horse PSEUDO-V telephone
'The horse V the telephone.'

c. Zam-tau tam maa-zai.
pillow PSEUDO-V horse
'The pillow V the horse.'

An important fact here is that the pseudo-verb (like *tam* in (61)) is compatible not only with two animate participants as in (61a), but also with an animate and an inanimate participants in both animate AGENT - inanimate PATIENT as in (61b) and inanimate AGENT - animate PATIENT as in (61c). They found that Cantonese-speaking children failed to comprehend sentences like (61a) and (61c), but they can understand (61b), where both word order and animacy support the AGENT.

Candan et al. (2012) is one of the few studies that has tested early acquisition of word order in Mandarin using the preferential looking paradigm. Their study focused on how English-, Turkish- and Mandarin-speaking children differ in manipulating word order as a cue in sentence comprehension. Test stimuli consisted in two simultaneous videos with theta-role reversal with existing verbs (e.g. *The horse is washing the bird* and *The bird is washing the horse*). Since they wanted to look solely at the weight of word order, the Turkish nouns were produced without case-marking, even though Turkish is a language with morphological case marking. The fact that the Turkish sentences are caseless in order to look more like English and Mandarin patterns may lead children to assume what they are hearing is something other than Turkish. I assume that the Mandarin sentences are well-formed, since there is no case marking anyway in Mandarin grammar. In fact, Candan et al. (2012) did not provide any examples of the test items in Mandarin in their paper. The results of Candan et al. (2012) indicated that English children showed early sensitivity to canonical word order at age 1.5, earlier than Turkish (2-year-olds) or Chinese children (almost age 3). As Chan et al. (2009), the authors also attributed the slow processing in canonical transitive sentences in Mandarin to the fact that in Mandarin both subject and object can be dropped, as well as the existence of varying word orders that makes canonical word order not the most reliable cue for sentence parsing. However, as pointed out by two anonymous reviewers from BUCLD-43 and EACL-10, English also has non-canonical structures. Therefore, it seems that Candan et al. (2012)'s argument that English has more rigid word order than Chinese does not stand up to scrutiny. The same explanation was given for the results on Turkish. Besides, for Mandarin, data collection was incomplete for the 1-year-old group and the measures of Candan et al. (2012) were not fully consistent, with gazing measures differing from number of switches of attention. Although Mandarin-speaking children were the less certain about matching sentences with scenes, they switched attention less frequently than Turkish-speaking children and did not differ from their English peers, which looked longer at the matching screen from very early on. A higher number of switches of attention is standardly interpreted as uncertainty in comprehension.

Interestingly, the experiment on another language with word order alternations and argument omissions, namely Japanese, reached the same conclusion: Omaki, Lassotta, Kobayashi, Rizzi, & Franck (2012) used the preferential-looking paradigm and followed the method of Franck et al. (2013) found that Japanese 19-month-olds fail to understand sentences with a canonical SOV order. Their corpus study revealed that 91% of the child-directed speech was uninformative to identify canonical word order. Only 12% of the sentences contained both subject and object arguments in Japanese. Moreover, it could be the case that Japanese children faced not only argument omissions but also case-marker omission in child-directed speech as reported in Matsuo, Kita, Shinya, Wood, & Naigles (2012), where they showed that case-marker is only available in 9% of the transitive sentences in parental utterances.

Mandarin behaves differently from previously studied languages in that it permits arguments to be null, presents no overt case marking, it is a topic prominent language but lacking agreement morphology (unlike other topic-prominent pro-drop languages, Chinese has no inflectional morphology associated with grammatical subjects/objects and topic phrases). Moreover, it presents lexical tones. For this reason, the study of Mandarin may shed light on early grammatical knowledge in languages with different grammatical features.

4.2 Research questions

As a result, whether Mandarin-speaking children are delayed in parsing canonical transitive SVO sentences, as suggested by Candan et al. (2012), or they can process them just as French or Hindi-Urdu children before 2 years of age, is to this day an open question. Some of the discrepancies between the studies presented so far could be partly explained as a methodological artefact. Chan et al. (2009) used judgement tasks. Although act-out tasks are easier than elicitation tasks, particularly for children with low MLUs, they are surprisingly difficult for very young children, since they require memory when planning an action (see Höhle, Bijeljac-Babic, Herold, Weissenborn, & Nazzi, 2009). Brandt-Kobele and Höhle (2010) combined different methodologies to test verb agreement in German-speaking children using eye-tracking techniques and a picture selection task. The same pairs of sentences were presented in both experiments. However, the eye-tracking data are interpreted as providing evidence that German-speaking children aged 3 to 4 are able to infer the number of the subject through verb infection, in contrast, children failed in the picture selection task, indicating a great impact of methodology on the comprehension found in children. Besides, research on the acquisition of syntax has reported asymmetries between comprehension and production, with a systematic advantage of the former (see Naigles, 2002 for a review). Although Candan et al. (2012) used the preferential looking paradigm (without an eye-tracker), a less cognitively demanding methodology, there was no data in their youngest group.

Thus, against this background, a first line of research consists in looking for new evidence for the acquisition of canonical word order in Mandarin Chinese. I address the question of word order by assessing the comprehension of canonical SVO order by Mandarin-speaking children at an earlier period, using eye-tracking measures. In particular, I replicate Franck et al. (2013) and combine the weird word order paradigm (which is often referred to in support of the lexical approaches) adapted to the IPLP (see section 4.1) with pseudo-verbs. This allows me to include my findings from Mandarin-speaking infants in a cross-linguistic comparison. Moreover, I conducted the same experiment with adults, while no results for adults are reported by Franck et al. (2013) or Gavarró et al. (2015).

4.3 Infant Mandarin

I report an experimental study using the Intermodal Preferential Looking Paradigm to examine canonical word order representations of 13 to 21-month-old children with a mean age of 17 months, as well as adult results[16].

[16]This section, coauthored with Julie Franck, Luigi Rizzi and Anna Gavarró, was originally published in: Zhu, J.T., Franck, J., Rizzi, L., & Gavarró, A. (2021). Do infants have abstract grammatical knowledge of word order at 17 months? Evidence from Mandarin Chinese. *Journal of Child Language*, first view, 1-20. (DOI: https://doi.org/10.1017/S0305000920000756).

4.3.1 Method

4.3.1.1 Participants

Seventeen typically-developing, Mandarin-speaking infants (7 boys, 10 girls) with a mean age of 17 months and 4 days (age range = 1;1.3-1;9.0, *SD* = 2.2) participated in my study. Seven additional infants participated in the study but they were not included in the results because of large errors in calibration (n = 4) or because of the infants' lack of eye tracking samples (n = 3). They were recruited in the Danan Community Health & Services Centre, China.

As a measure of the infants' linguistic development, their vocabulary was assessed using the Mandarin version of the Communicative Development Inventory (CDI, Hao, Shu, Xing, & Li, 2008), which consists of two checklists: an infant checklist (used for infants between 12 and 16 months of age) and a toddler checklist (used for children between 17 and 30 months of age). Both checklists included the animals' names used in my study. Following Hao et al. (2008), for the toddler list, parents were only asked to indicate whether their children had ever said the word, as is done in the English CDI (Fenson et al., 1994), and so no comprehension scores are available.

For my study, infants from 13 to 16 months (the younger group, n = 8) achieved a mean score of production of 5 words (Standard Deviations *SD* = 5.5, range from 0 to 11) and a mean score of comprehension of 25 words (*SD* = 14.2, range from 9 to 36). Infants from 17 to 21 months (the elder group, n = 9) achieved a mean score of production of 43 words (*SD* = 31.4, range from 0 to 102). The summary of their scores is shown in table 17.

Table 17: Infants' vocabulary.

Months	Comprehension (words) Mean	Range	Production (words) Mean	Range
13-16	25 ($SD = 14.2$)	9-36	5 ($SD = 5.5$)	0-11
17-21	-	-	43 ($SD = 31.4$)	0-102

4.3.1.2 Materials and design

Franck et al. (2013) resorted to the WWO paradigm, and children were confronted with both grammatical NP-V-NP and ungrammatical NP-NP-V sequences, three sentences in each condition. NP-NP-V is considered ungrammatical, because, first, it is used in neutral contexts and with neutral intonation, whereas it is only admitted in contexts that license contrastive focus, and involves a special focal intonation (Tsai, 2008). When the object is animate, *ba* is obligatorily required in neutral contexts and with neutral intonation (Van Bergen, 2006). Second, in a recent grammaticality judgement task (Yu & Tamaoka, 2018), the animate-animate verb sentences without *ba* were judged of very low acceptability, and mostly regarded as uninterpretable among native speakers (2.9% of acceptability in younger group with mean age of 25.8 years old and 10.3% of acceptability in older group with mean age of 56.6 years old). Third, bare NP-NP-V sentence are very rare in child-directed speech (less than 1%, according to the study of Yeh, 2015).

(62a) and (62b) correspond to the grammatical and ungrammatical sentences used in the experiment. In Chinese, aspectual information is systematically expressed; besides, the perfective marker *le* to mark the end of the action is used far more frequently than other markers in early speech (Erbaugh, 1982). For that reason, the perfective aspect *le* was selected to describe the scene. The six test items (three in each condition) are listed in Appendix A.3.

(62) a. Xiao-gou chei le xiao-lv.
little-dog PSEUDO-V PERF little-donkey
'The puppy V-ed the little donkey.'

b. *Xiao-lv xiao-gou nui le.
little-donkey little-dog PSEUDO-V PERF
'The little donkey the puppy V-ed.'

The two monosyllabic words *nuí* 'to put a crown on someone's head' and *chéi* 'to put someone's head under a net' were created for the purpose of this study. Verbs in the phonological neighborhood of the two pseudo-verbs showed a similar distribution of transitivity. By phonological neighborhood, following Luce and Pisoni (1998), I refer to the number of words that differ in phonetic structure from a target word by only a single unit, that can be either substituted, deleted, or added. Statistics computed on the number of verbs used in the Han's corpus (2009) showed that 61.3% of the verbs in the phonological neighborhood of *nuí* were transitive, while the distribution was 60% for *chéi*. *Chéi* was used in the NP-V-NP condition, whereas *nuí* was used in the NP-NP-V condition.

To verify that my pseudo-verbs followed the phonological pattern and phonotactic constraints of Mandarin verbs, I asked ten adult Chinese speakers to judge if each verb (embedded in a sentence) sounded familiar and whether they knew its meaning. The judgement was based on a binary scale (yes/no). All ten participants said the verbs sounded familiar, but could not assign any meaning to it.

Mandarin being a tone language, the pseudo-verbs used in the test presented a high tone, and lexical tone interacts with sentential intonation. I compare the pitch movements and pitch range expansion in three of the test sentences, with lexical tone kept constant, using Praat (Boersma & Weenink, 2005). The intonational pattern of test SVO sentences is illustrated in Fig. 7. In Fig. 8, I show the intonational pattern of the ungrammatical SOV sentences, which was the same as the intonational pattern of their grammatical (*ba* construction) counterparts in Fig. 9, as both were rising-falling-rising, with pitch accent on the first NP. Thus, the ungrammaticality of the SOV sequences in my experiment stemmed only from word order, rather than the intonational pattern imposed on the sequence.

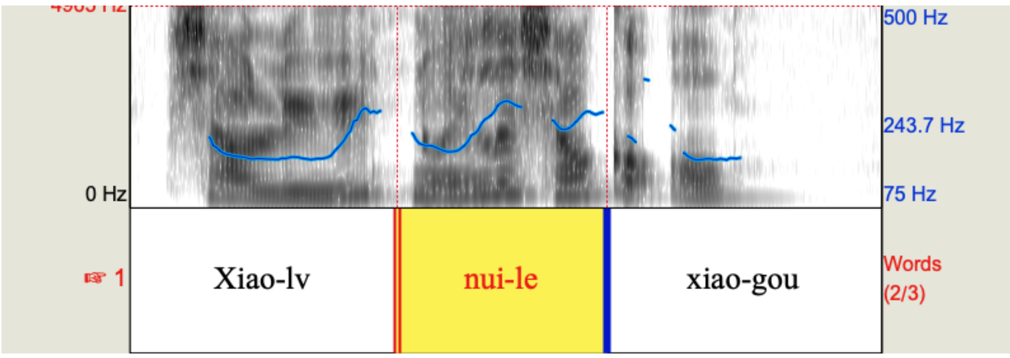


Figure 7: Intonational pattern of the grammatical SVO sentence *xiao-lv nui-le xiao-gou* 'The little donkey V-ed the puppy'.

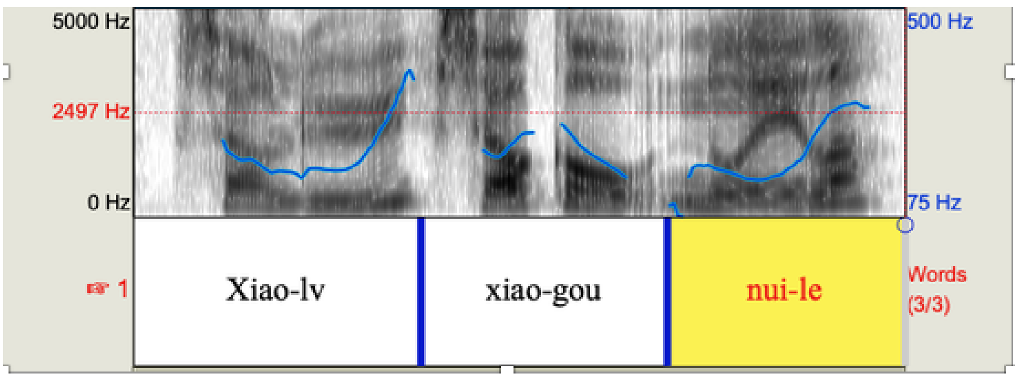


Figure 8: Intonational pattern of the ungrammatical SOV sentence *xiao-lv xiao-gou nui-le*.

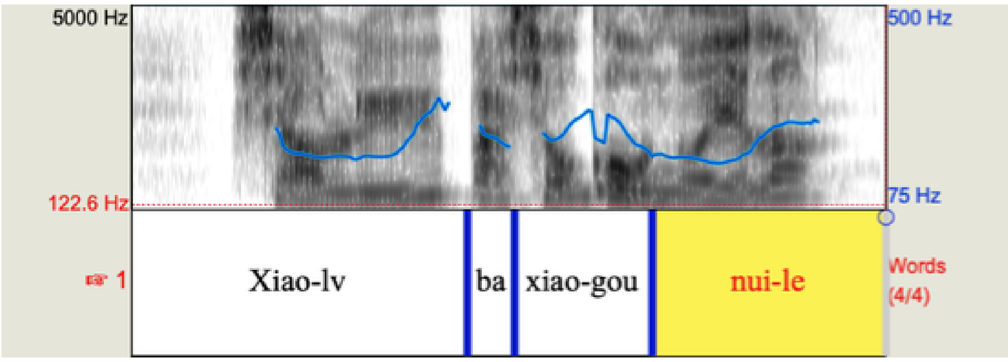


Figure 9: Intonational pattern of the grammatical S*ba*OV sentence *xiao-lv ba xiao-gou nui-le.*

Visual stimuli were the same as in Franck et al. (2013), and the characters included puppets of a dog, a donkey, a lion, a horse, a cow and a sheep. I added the adjective *xiao* 'little' to adapt the video culturally. The sound track was pre-recorded in normal declarative intonation by a Mandarin female native speaker (30yr) in a sound-proof booth, recorded and digitized at 44.1 KHz, mono-channel. Utterances were chopped using Praat (Boersma & Weenink, 2005) to make sure that all repetitions were the same, and videos were edited with Adobe Premier Pro CC 2017 (v. 11.0.2).

In the experimental session, for each sentence, the participants were presented with two simultaneous videos, one video shows the action carried out transitively with the first NP as AGENT and the second NP as PATIENT (e.g. the cow puts a crown on the lion's head), the other video illustrates the same action carried out reflexively with both NPs as AGENTS (e.g. the cow and the lion each put a crown on their own head). The items were presented in random order with the presentation of the transitive and reflexive event counterbalanced across the left and right sides of the screen and across the grammatical and ungrammatical conditions. An example of the materials used in the experiment is given in Figure 10.

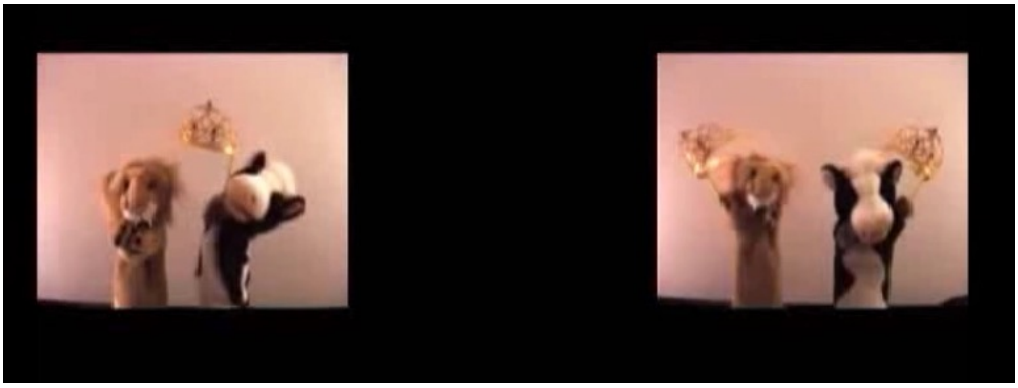


Figure 10: Test material, experiment 1.

4.3.1.3 Procedure

The eye-tracker used was a Tobii Pro X3-120 and Tobii Studio^{TM} (Version 3.4.8) was used as platform for the recording and analysis of eye gaze data. Note that I collected the data using an eye-tracker, which measures gaze direction objectively, as opposed to video recordings of subject's eye movements, as in earlier work.

The video stimuli were projected from a laptop. Each child sat on his or her caregiver's lap approximately 60 cm from the computer screen during the whole length of the experiment. The caregivers were asked to close their eyes and listen to music played through headphones during the test trials so as not to guide their children toward any of the videos. Testing occurred in a bright small room (300-350 Lux, Temperature 18-25^{o}C), which did not have windows.

The experimental session started with the procedure of eye calibration. An operator-controlled calibration was run, which consisted of a yellow chick presented at the four corners and the centre of the screen.The chick was accompanied by a sound. When the operator judged the participant to have finished looking at the chick, a button was pressed. Details of the calibration stimuli are given in Hessels, Andersson, Hooge, Nyström, & Kemner (2015). The operator judged the calibration output from the Tobii Pro, after which a decision was made to accept the calibration or re-calibrate.

Once the child's gaze director vector was calibrated, I proceeded to the training session. First the participants went through a character-introduction phase; all the puppets were presented once (e.g. *Bao-bao kuai kan, shei zai na-li? O, shi xiao-lv* 'Look, who's there? It's the little donkey'), while the other screen remained blank (6s) (see Fig.11). To keep the interest to the maximum, I varied the carrier sentence. There were three possible carrier sentences (*Bao-bao kuai kan* 'Baby, look!'/*Kan dao le ma?* 'Did you see'/ *shei zai na-li* 'Who's there?'), counterbalanced across trials.

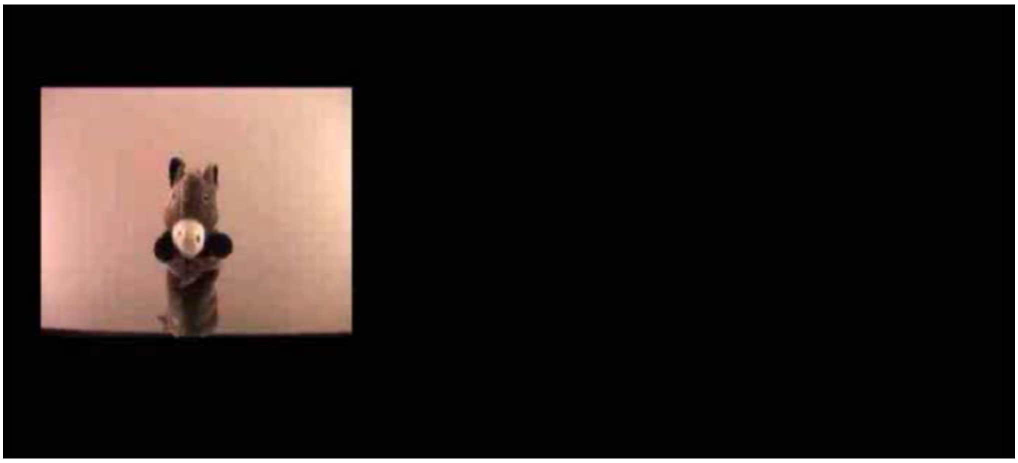


Figure 11: The character-introduction video clip, experiment 1.

Next, the participants were introduced to the simultaneous presentation (see Fig. 12), which showed two different animals at the same time while the recorded audio asked them to find one of them (e.g. *Bao-bao* *kuai kan, kan-dao xiao-lv le ma? Xiao-lv zai na-li ya?* 'Look, do you see the little donkey? Where is the little donkey?').

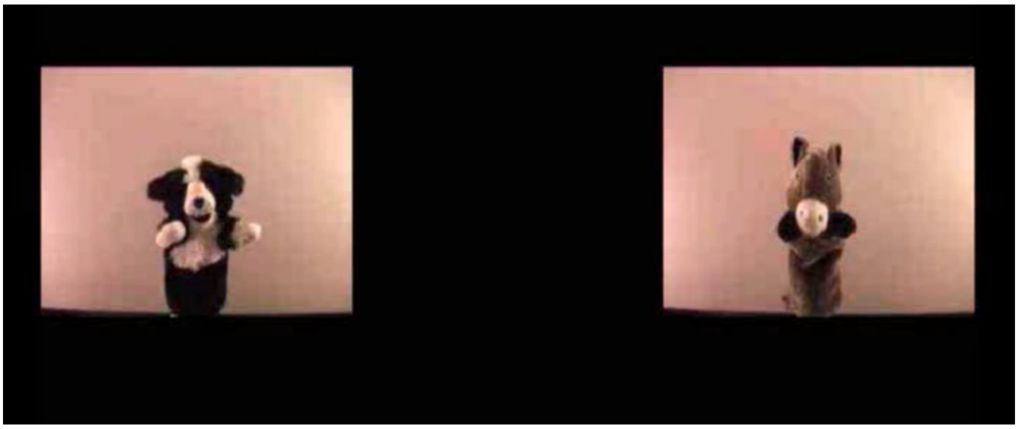


Figure 12: The character-identification video clip, experiment 1.

Finally, the participants saw the novel actions used (see Fig. 13); most importantly, novel actions were presented in neutral frames without the use of the novel verbs, paired with sentences like *Kan, fa-sheng le shen-me?* 'Look, what happened?' such that later understanding of the test sentences cannot be attributed to lexical learning during the training phase. Ambridge and Lieven (2011) questioned the role of the training phase. They argued that infants may learn some relevant syntactic information that would help them parse the test sentences (see also Franck et al., 2013 for discussion). However, in my study, no linguistic information, lexical or structural, was provided in the training phase.

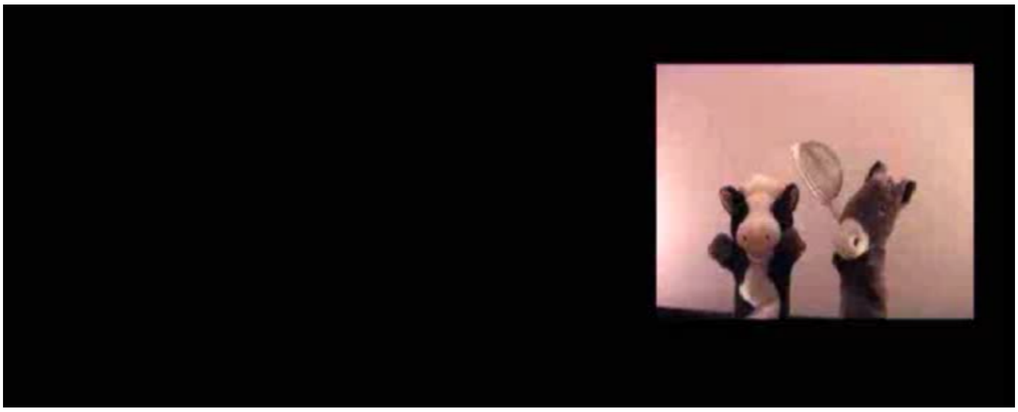


Figure 13: The video clip introducing the novel action used in the experiment 1.

After the training session and a short transition cartoon, the experimental session started (see Fig. 14 and Fig. 15). A blank screen (2s) appeared between experimental items (six in total), and after item 3, 4 and item 5 a clip of a Teletubbies landscape was shown to keep the child's attention. All videos started with a sentence to draw his/her attention (e.g. *Baobao, kai kan, fa-sheng le shen-me* 'Look, what happened?') as baseline, and then the experimental sentences were played three times. The recording of gazing time took place in four windows: the baseline and three consecutive exposures to the target sentence starting at 5, 10 and 15 seconds. The whole session lasted between 10 and 15 minutes. After the test session, the experimenter asked the infants' caregivers to fill out the Chinese version of CDI (Hao et al., 2008).

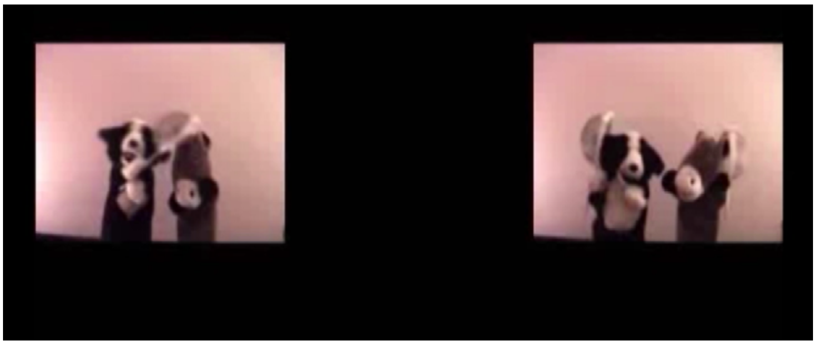


Figure 14: Grammatical sentence with *chéi*

Figure 15: Ungrammatical sentence with *nuí*

4.3.1.4 Predictions

The prediction is that if adult-like abstract syntactic knowledge is available to children from early on, they will show a significant preference for the causative scene upon presentation of NP-V-NP sequences. Based on the findings in the literature, I predict lack of preference in the ungrammatical NP-NP-V sequences. In contrast, if word order is encoded as a lexical property, children will show random behavior in all conditions, as pseudo-verbs were used.

4.3.1.5 Data treatment and analyses

Following Franck et al. (2013), only infants whose detected signal was more than 55% were taken into account. The number of participants analyzed was 17.

To provide an overview of the eye gaze pattern, I used statistical analyses involving bivaried paired Student t tests for means' comparisons and linear mixed-effects models (LMMs) for total fixation duration using the R software package (i.e., lme4, Bates et al., 2014), version 3.5.2 (R Development Core Team, 2015). In the full model, the fixed effects included Scene (Transitive vs. Reflexive) and Condition (Grammatical vs. Ungrammatical) and their interactions; the random effects included items and participants, where both their intercepts and slopes were allowed to vary among all the fixed effects (Baayen et al., 2008). I used the Wald test to compute p-values for each fixed effect.

Proportion of looking times to the transitive scene was also calculated. Since the response variable was proportions, which follow a multinomial distribution (i.e. between 0 and 1) rather than a normal distribution, traditional LMMs or ANOVA methods based on normal distribution cannot be directly used (Barr, 2008). As a result, I applied Generalised linear mixed models (GLMMs), which do not require the assumption of normal distribution (Bolker et al., 2009). In the full model, I treated RoIs (Baseline, Sentence 1, Sentence 2, Sentence 3), Condition (Grammatical vs. Ungrammatical) and their interactions as fixed factors, with items and participants as random effects. If the interaction were significant, then I explored each RoI separately with Condition as fixed factor and proportion of looks to the transitive scene as dependent measure, items and participants were included as random effects.

Finally, to explore the effect of age and vocabulary on proportions of fixations to the transitive video, I applied GLMMs on the RoIs that showed a significant effect of Condition, with the proportion of looks to the transitive video as dependent variable and Condition, Vocabulary (as continuous variable) and Age (as continuous variable) as factors.

4.3.2 Results

First, the mean looking times to each of the videos (transitive vs. reflexive) in each of the four windows were considered: the baseline window and the three consecutive windows corresponding to first, second and third exposure to the experimental sentence are reported in table 18. The results are expressed in milliseconds (ms).

Table 18: Mean looking times (in ms, standard deviations in parentheses) across the four critical RoIs, infants.
*p <.05 (in bold)

	Grammatical		Ungrammatical	
	Transitive	Reflexive	Transitive	Reflexive
Baseline (0-4s)	1438 (619)	1098 (689)	1394 (818)	1138 (439)
Sentence 1 (5-9s)	**1386 (665)***	**928 (724)***	1198 (779)	1072 (569)
Sentence 2 (10-14s)	**1266 (903)***	**793 (831)***	974 (680)	1041 (1012)
Sentence 3 (15-19s)	931 (772)	852 (662)	1153 (1021)	876 (779)

Paired t tests were conducted on meaning looking times. Infants looked significantly longer at the transitive video than at the reflexive video only in the grammatical condition, during the first (t(16) = 2.16, p = .046, Cohen's d = .52) and the second presentation of the test sentence (t(16) = 2.12, p = .049, Cohen's d = .51). No significant difference was found in the baseline window nor in any of the windows of the ungrammatical NP-NP-V condition nor in the third presentation of grammatical NP-V-NP condition.

For multiple comparisons, I run two LMMs separately for each RoI based on total fixation duration: one in the grammatical condition, and one in the ungrammatical condition. The model used looking time in the 0- to 4s window as baseline against which the evolution of looking preferences in the following windows were estimated.

The successful model was only found in the grammatical condition. More specifically, there was a significant interaction between the first presentation and the looking time to the transitive window (β = .12, t = 6.03, p = .046), the same can be found between the second presentation and the looking time to the causative window (β = .13, t = 11.39, p = .046) with positive coefficient, which shows that infants looked significantly more at the transitive scene after the first and the second presentation of the sentence when hearing the grammatical SVO sentences. This confirms the significant preference for the causative scene in the grammatical condition and null results in the ungrammatical condition reported in table 18.

An example of the heat maps for the grammatical sentence (62a) for all infants is illustrated in Fig. 16. Red indicates the highest number of fixations or the longest time, and green the least; as can be observed, they fixed their gaze longer in the transitive action as shown by the thicker red shade for the intensity of gaze, while this intensity effect was fluctuating in the ungrammatical sentences (62b) as can be observed in Fig. 17.

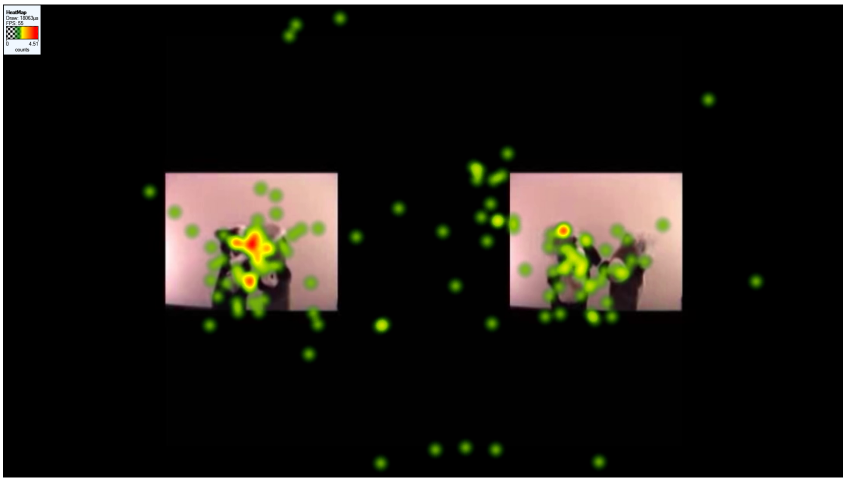


Figure 16: Heat map for the grammatical sentence *Xiao-gou chei le xiao-lv* ‘The puppy V-ed the little donkey’. Red indicates the highest number of fixations or the longest time, and green the least. The left video represents the transitive event and the right video represents the reflexive event.

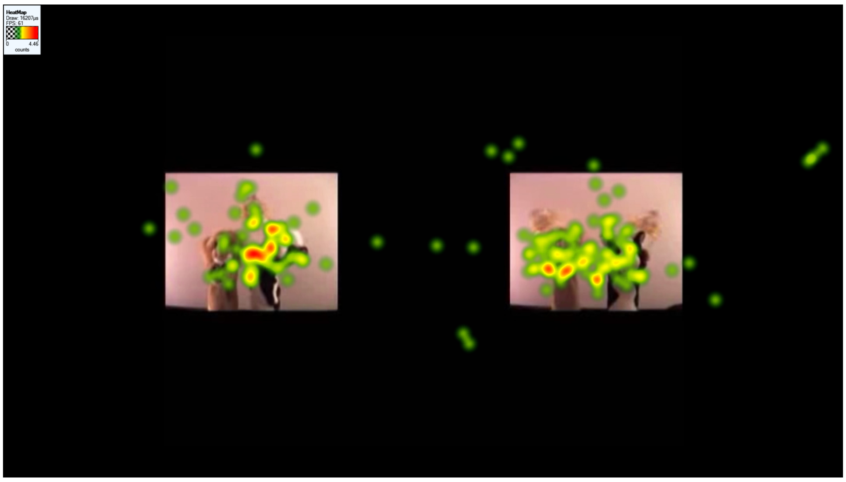


Figure 17: Heat map for the ungrammatical sentence *Xiao-liu shi-zi nui le* ‘The calf the lion V-ed’. The left video represents the transitive event and the right video represents the reflexive event.

I also calculated the proportion of looking time to the transitive video (calculated over total looking time to the transitive and reflexive videos) in the four critical RoIs. As shown in Fig.18, preference for the transitive over the reflexive action in the grammatical condition emerges in the first presentation of the sentence (mean = .62), peaking at the next window after the second presentation of the test sentence (mean = .64), and then decreases in the last window (mean = .54).

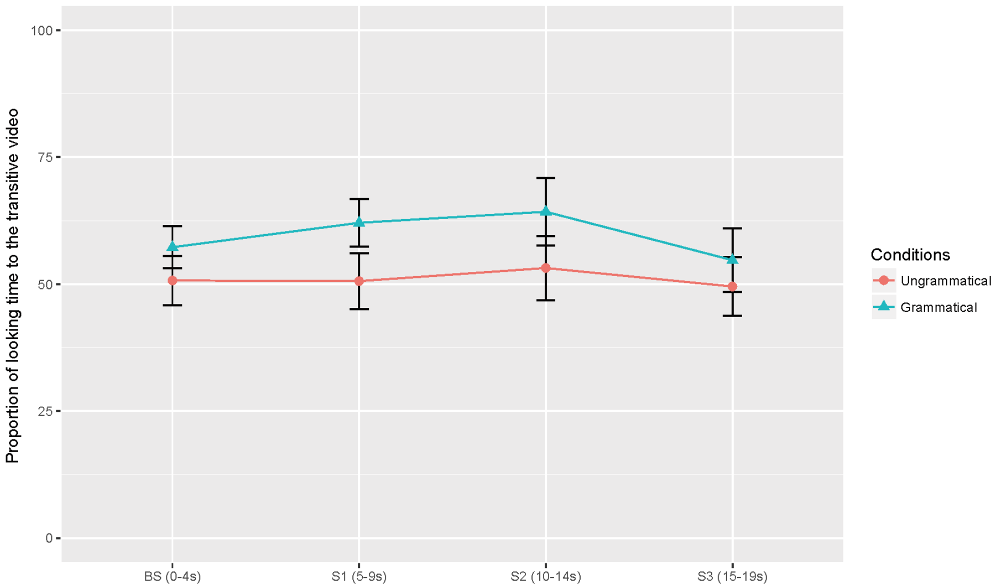


Figure 18: Proportion of looking time to the transitive video in the four critical RoIs, infants.

One main issue concerns the fact that a preference for the transitive scene can already be seen in the grammatical condition in the baseline (see Figure 18). What we really want to know is whether this preference increases in the next RoIs in the grammatical condition more than in the ungrammatical ones. To do this, I ran GLMMs with the proportion of looking times to the transitive action as dependent variable and RoI (Baseline, Sentence 1, Sentence 2, Sentence 3) and Condition (Grammatical, Ungrammatical) as factors.

The results showed a significant interaction between RoI and Condition (z = .46, SE = .11, p = .045), which allowed us to further explore the effect of Condition in each RoI. I found a significant effect of Condition after the first presentation of the sentence (β = .29, t =1.43, p = .016) and the second presentation (β = .11, t = 1.49, p = .027), which means that infants showed an increased preference to look at the transitive video compared to the reflexive one when they heard a grammatical sentence compared to when they heard an ungrammatical one. No effect of Condition was found in the baseline window (β = .15, t = .84, p = .41) nor after the third presentation of the sentence (β =- .07, t =- .35, p = .72).

Finally, I explored the main effect of age and vocabulary. GLMMs ran on the two RoIs that showed a significant effect of Condition (i.e. S1 and S2 together), with the proportion of looks to the transitive video as dependent variable and Condition, Vocabulary and Age as factors, showed no effect of Age (β = -.01, t = -.44, p = .66), no main effect of Vocabulary (β =.016, t = 1.74, p = .34), and critically no interaction between Vocabulary and Condition (β = -.016, t = -1.25, p = .22), nor between Age and Condition (β = -.021, t = -.69, p = .49). This indicates that neither vocabulary nor age modulated the effect of grammaticality. The three-way interaction was not significant either (β =.0009, t = 1.2, p = .24). This confirms again that the increased preference found for the transitive video over the reflexive video when a grammatical sentence is presented is independent from age and vocabulary.

4.4 Adult Mandarin

The same experiment was conducted to discover adults' comprehension of canonical word orders with pseudo-verbs in Mandarin, while in Franck et al. (2013) no results for adults are reported.

4.4.1 Method

4.4.1.1 Participants

Eighteen native Mandarin-speaking adults (age range = 24-53, mean age = 29, SD = 7.7) participated in my study. They were recruited in Guiyang, China and Barcelona.

4.4.1.2 Materials, procedure, data treatment and analyses

The materials and procedure used were the same as those for infants. I adopted the same analysis for adults' data as had been adopted for the infant data. For all the adults tested, the detected signal was more than 75%.

4.4.2 Results

The mean looking times to each of the scenes in the four RoIs for adults can be found in Table 19.

Table 19: Mean looking times (in ms, *SD* in parentheses) across the four critical RoIs, adults.
***p <.001 (in bold)

	Grammatical		Ungrammatical	
	Transitive	Reflexive	Transitive	Reflexive
Baseline (0-4s)	1765 (511)	1437 (503)	1833 (533)	1567 (556)
Sentence 1 (5-9s)	**2585 (617)*** **	**909 (397)*** **	1711 (613)	1829 (662)
Sentence 2 (10-14s)	**3140 (967)*** **	**325 (504)*** **	1626 (1070)	1639 (963)
Sentence 3 (15-19s)	**3267 (720)*** **	**379 (451)*** **	1730 (740)	1790 (711)

In the case of adults, except for the baseline window, adults looked significantly longer at the transitive video in the grammatical condition during the first (t(17) = 9.66, p < .001, Cohens d = 2.27), the second (t(17) = 9.13, p < .001, Cohens d = 2.15) and the third presentation of the test sentence (t(17) = 12.49, p < .001, Cohens d = 2.95). No significant difference was found in the baseline window nor in any of the windows of the ungrammatical NP-NP-V condition (all p > .05).

The mixed model on total looking times only showed an effect in the grammatical condition: there was a significant main effect of Scene (β = .40, t = 1.96, p = .037). Besides, the positive coefficient between the scene and the looking time indicated there was an increase of looking times to the causative video over time (see table 20).

Table 20: Fixed effects from the best-fitting model of looking times in the grammatical condition, adults.
***p <.001 (in bold)

Fixed effects	Estimate	*SE*	*t* value
(Intercept)	1.37	0.12	**11.56*****
Scene (Transitive): RoIs (S1)	1.28	0.21	**6.03*****
Scene (Transitive): RoIs (S2)	2.42	0.21	**11.39*****
Scene (Transitive): RoIs (S3)	2.48	0.21	**11.56*****

Formula in R: Looking times ∼ Scene*RoIs+(1+Scene+RoIs|Subject)+(1+Scene+RoIs|Item)

The proportion of looking time to the transitive video in each RoI is illustrated in Figure 19.

The generalized linear model with the proportion of looking times to the transitive action as dependent variable and RoI and Condition as factors showed a significant interaction between the two factors (z = 1.92, SE = .98, p = .04). Thus, I explored the effect of Condition in each RoI separately. With the generalized model I found a significant main effect of Condition during the first presentation (β = .86, t =5.64, p < .001), the second presentation (β = 1.59, t =5.54, p < .001) and

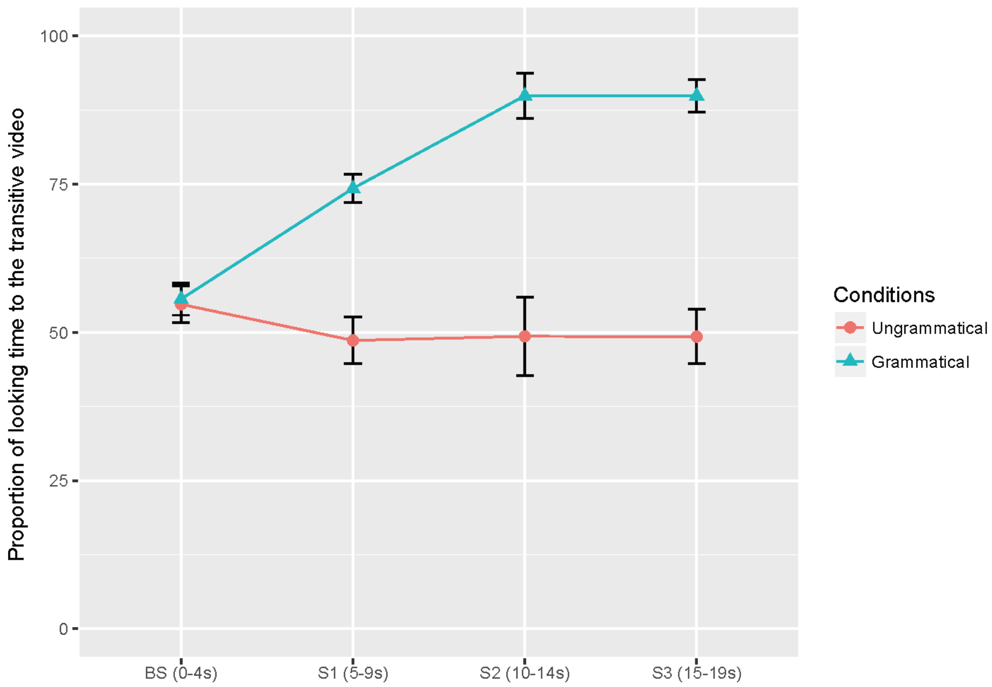


Figure 19: Proportion of looking time to the transitive video in the four critical RoIs, adults.

the third presentation of the sentence (β =1.56, t =7.82, p < .001), showing that the preference for the transitive video increased when a grammatical sentence is presented. No effect of Condition was found in the baseline window (β = .03, t = .23, p = .81).

4.5 Discussion

I tested the comprehension of canonical transitive NP-V-NP sentences in very young learners of Mandarin combining the WWO paradigm with ungrammatical NP-NP-V sequences with the preferential looking paradigm using eye-tracking techniques (as in Franck et al., 2013 and Gavarró et al., 2015). My results indicate that infants acquiring Mandarin show a preference for the transitive scene when they encounter the grammatical transitive NP-V-NP frame with novel verbs, but that does not happen when they hear ungrammatical NP-NP-V sequences. Besides, the results for adults are very similar to those for infants: with grammatical sequences, adults direct their gaze towards the transitive video, with ungrammatical sequences they direct they gaze randomly across the two videos. The only difference between adults and infants is that adults maintain attention with a grammatical sequence on the transitive video until the last presentation of the sentence.

In contrast, some work conducted in our lab showed that Finnish adults obtained above-chance-level performance when hearing ungrammatical SOV sentences (as in (64)) in an experiment with the same design (Keidel, Leela, Zhu, & Kunnari, 2021).

(63) Leijona täkee hevosen.
lion-NOM PSEUDO-V horse-ACC
'The lion V-ed the horse.'

(64) *Lammas hevosen raistaa.
sheep-NOM horse-ACC PSEUDO-V
'The sheep the horse V-ed.'

In the French study, adult French speakers also directed their gaze towards the transitive video when when hearing ungrammatical SOV sentences (Franck, p.c.). The fact that, using the same experiment, Mandarin adults performed at chance level upon encountering ungrammatical sentences while adults in Finnish did not, is still an unresolved issue. One may think that the same participants heard both grammatical and ungrammatical sentences within the same experimental session, a principle of contrast applied (Clark, 1987), namely, adults may expect different forms to contrast in meaning, that then would map verbs appearing in the NP-V-NP frame to a causative event, and the verb appearing in the other frame (i.e. NP-NP-V sequences) to an non-causative event. But that was not the case, in Finnish, as adults looked at the transitive scene even when hearing ungrammatical sentences. One explanation could be that word order is more restrictive in Mandarin than in Finnish, as all mathematically possible word order permutations can be found in Finnish (Vilkuna, 1995), while only four of them can be found in Mandarin. Besides, in a very recent study, Brattico (2021) examined word order variations in basic finite clauses in Finnish and their grammatical estimations. He found that SOV in Finnish was judged only as marginal while SOV in English was judged as ungrammatical.

Another fact regarding the baseline, the fact that infants, as adults, looked numerically longer at the transitive scene in the baseline, has been observed in previous studies by Naigles (1996), in which 2-year-old children looked longer at transitive actions than contact actions (in which one character merely touches the other) in the Frameless condition (e.g. *Look! Sebbing!*). The slight preference for the transitive action was nevertheless taken into account in my analysis.

The preference observed for infants in the grammatical condition cannot be explained by the lexical approaches, since all sentences included pseudo-verbs. Neither comprehension of the grammatical sequence, nor the difference in performance between grammatical and ungrammatical sequences is predicted by the lexical approach. A recent study by Hsu, Rispoli, & Hadley (2019) also provides counter-evidence for lexical approaches although for a different phenomenon: despite the fact that instances of Mandarin resultative verb compounds -*de* (as in (65)) were four times less frequent than those of -*bu* (as in (66)) in naturalistic speech, there was no difference in children's performance in elicitation tasks at age 3.

(65) chui de po
blow can break
'has the ability to blow (into something) and make it burst'

(66) chui bu po
blow no broke
'does not have the ability to blow (into something) and make it burst'

In addition, provided with appropriate discourse contexts, children did extremely well producing resultative verb compounds (as *kan-bu-dao* 'not able to see' and *kan-de-dao* 'be able to see') even though resultative verb compounding occurred in less than 0.7% of parental input, which also runs against the frequency-based approach of language acquisition of the variational model (Yang, 2002, 2004).

The performance attested in the results here also runs against the predictions of a grammar-based approach which claims that infants follow an AGENT-first strategy in their parsing (Bates & MacWhinney, 1982; Bever, 1970; Lidz et al., 2001). The results from Gertner et al. (2006) could indeed be interpreted as such, given that the two videos used illustrated transitive actions with reversed theta-roles and the test sentences were all subject first. If infants had proceeded in that way in my experiment, they would have performed identically with NP-NP-V and NP-V-NP sequences, since the two videos illustrated the first NP as the AGENT. Even if an AGENT-first strategy exists, my results show that it cannot override grammaticality, that is, it cannot be used to assign an interpretation to an ungrammatical sentence. This observation was already made by Franck et al. (2013) and Gavarró et al. (2015).

The grammatical NP-V-NP sentences (like (62a)) include known nouns and an unknown verb, and the correct interpretation of such structure implies that young children can use the arguments in a sentence to infer the syntactic structure and take the unknown word to be the verb (i.e. syntactic bootstrapping). This implies that as soon as infants comprehend some words (at around 12 months according to Oviatt, 1980, or even before), they may start accessing the meaning of the whole sentence; likewise the results provide evidence that children can infer the subcategorization frame of a novel verb based on the syntactic structure, namely, when they hear a verb describing a two-argument event in a target NP-V-NP manner, they infer that the verb has a transitive meaning, whereas if the two arguments appear in a NP-NP-V frame they fail to parse the sentence, in particular, they do not identify the immediately preverbal NP as the object. Besides, recall that the ungrammaticality of the test items in my experiment arises from word order alone, since they have been produced just as their grammatical counterparts, which shows again that the behavior observed in NP-NP-V sequences cannot be taken as a sign that infants lack syntactic competence; as with adults, I attribute the infants' behavior to sensitivity to the ungrammatical sequence.

Since a novel verb was used, the ungrammatical sentences could also be regarded as sentences with a null object with a coordinated NP subject. This is not an option for adults here, since the object can only be omitted given the proper discourse context and the reference should be picked up from the previous discourse. There is the possibility that if infants treat the verb as intransitive (as there is no NP following the verb), Mandarin infants might look more at the reflexive events when they hear the NP-NP-V sequences since there is a bias to map intransitive frames to reflexive events (Chang, Dell, & Bock, 2006). This is not what the results show.

Moreover, in the Hindi-Urdu experiment (Gavarró et al., 2015), infants looked at the causative event when they heard the NP-NP-V (i.e. SOV) sequence, but no preference was shown when they heard the V-NP-NP (i.e. VSO) sequence, although the two NPs could be taken to be conjoined again. For those reasons I think that children are not being misled into interpreting NP-NP sequences as coordinations (this interpretation would furthermore require the omission of the coordination).

Compared to other experiments in the same paradigm, the results of my study are in line with the original French experiment (French being an SVO language with little presence of null arguments), as well as the Hindi-Urdu experiment (Hindi-Urdu being an SOV language with generalized pro-drop). The only difference in the results is due to the grammatical difference between SVO and SOV languages: French and Mandarin-speaking infants increased significantly their fixations to the transitive video when they heard the grammatical NP-V-NP sequences compared to the NP-NP-V word order, ungrammatical in these languages. By contrast, the NP-NP-V order gave rise to a looking preference for the target transitive event in infants acquiring Hindi-Urdu, an SOV language. This shows that infants are sensitive to the specific syntactic structures of the languages they are exposed to.

Due to the length of each experimental item, comparison between the exact timing of effects among the three languages is not really possible: in the French experiment, the 20-second video was split in 5 windows, while both in Hindi-Urdu and the present experiment there were 4 windows of analysis. However, I can still make some observations. Preference for the transitive over the reflexive action appears in window 8-12s, and the effect was significant in the 12-16s in the case of French infants, and 6-10s in the case of Hindi-Urdu, while in Mandarin the effect was significant in 5-9s. Hindi-Urdu was the only language in which the effect persisted until the last 16-20s window, while in Mandarin and French the effects disappeared in the last window, which could be due to tiredness, since, at least in Mandarin, infants were younger[17].

[17] Why does the effect appear late in French compared to Hindi-Urdu or Mandarin? A closer look at the test sentences would give us a perhaps speculative answer (see (i)). In the French experiment, the pseudo-verbs in the test sentences did not involve any aspectual o functional information (ia), while in both Mandarin (ib) and Hindi-Urdu (ic), the verbs contained perfective aspect (like *le* in Mandarin, and *−(y)aa* in Hindi-Urdu, besides additional case markers in Hindi-Urdu), which could help infants identify the verb, just as they use determiners to bootstrap nouns by 12 months (Kedar, Casasola, Lust, & Parmet, 2017).

(i) a. Le lion poune le cheval.
DET lion PSEUDO-V DET horse
'The lion V-ed the horse.' (French)

b. Shi-zi chei le xiao-ma.
lion PSEUDO-V PERF little-horse
'The lion V-ed the foal.' (Mandarin)

c. Sher-ne ghode-ko khalaa-yaa.
lion-ERG horse-ACC PSEUDO-V-PERF
'The lion V-ed the horse.' (Hindi-Urdu)

Previous studies have found that infants from 12-16 months are able to use function words to categorize novel words (Höhle, Weissenborn, Klefer, Schulz, & Schmitz, 2004; Zhang, Shi, & Li, 2015) and 18-month-olds can use function words to recognize verbs (Cauvet et al., 2014). Still, the presence of overt functional elements is not essential, as witnessed by the original results from French, where infants at 19 months were able to parse the grammatical sentences with no overt functional element.

In my study, the vocabulary scores of the infants exposed to Mandarin were lower than those of the French infants (for the infants exposed to French the mean was 87, and the range was 8–389), while no results for vocabulary were given in the study on Hindi-Urdu. In Franck et al. (2013), children's lexical knowledge failed to predict individual preferences for matching video. Moreover, age was not a predictor of comprehension either. Mandarin children here were younger by almost two months (17.4 vs. 19) when compared with infants in Franck et al. (2013). Thus the findings of Franck et al. (2013) are now replicated with infants younger than in that previous study.

It would seem, then, that Mandarin, French and Hindi-Urdu would pattern alike, and therefore there would be no grounds to establish a cross-linguistic difference in the emergence of early syntax, as far as fundamental word order properties are concerned, at least for languages like Mandarin, French and Hindi-Urdu, despite differences in word order, case marking, and availability of pro-drop.

Thus, my results contrast with those from Candan et al. (2012) for Mandarin, who only found evidence for word order acquisition at almost age 3. Besides, the results are also at odds with Chan et al. (2009), who found evidence for the acquisition of canonical word order only at 3;6 with an act-out task. Candan et al. (2012) used the preferential looking paradigm as well, and I hypothesize that the different results may due to the following four factors. First, actually there was no data for 1-year-olds in Candan et al.'s work (2012), their data was only available for the 2-year-old and 3-year-old groups in the case of Mandarin. Second, the study by Yang, Shi, & Xu (2018) reveals that the perfective marker *le* did have an immediate effect on 30-month-old Mandarin-speaking children's looking behavior: as soon as they heard *le*, they looked at the scene in which the event began and terminated, while they showed latency in looking at scenes matching sentences with the imperfective marker *zhe*, which describes an on-going, progressive event. This result comes to show that the presence of *le* facilitates sentence processing, thus children may have benefited from the aspect marker *le*, which in Candan et al.'s items was either absent or may be replaced by imperfective *zhe* following their English and Turkish items[18], which does not have an immediate facilitating effect when compared to *le* (Yang et al., 2018). In the longitudinal study of Erbaugh (1982), *le* appeared earlier than *zhe* in child production: *zhe* started to appear at 2;3 while *le* was already used productively by two years of age; and was used far more frequently than *zhe* in early speech: children produced up to 2300 - *le* in their early speech in contrast to the total of 50 - *zhe* in the corpus of Erbaugh (1982). These studies converge in showing that *le* is acquired earlier than *zhe*, possibly due to its higher frequency in the input. Third, children from Candan et al. (2012) were recruited in Taiwan, where apart from Mandarin, children could be also exposed to the Taiwanese Southern Min dialect, which is a strongly OV language (Huang & Roberts, 2017). This may have affected their performance when confronted with NP-V-NP sequences, as they could be interpreted as the target SVO or the distractor OVS influenced by Taiwanese Southern Min dialect. Finally, in my study, the target video depicted a causative action, while another depicted a reflexive one, with no theta-role reversal, while distractors with theta-role reversal were used in Candan et al. (2012). As pointed out by Yang

[18]In fact, Candan et al. (2012) did not provide examples of the test items in Mandarin

et al. (2018), reversibility of NPs may have complicated the processing task. However, this last argument will not be maintained, as shown in chapter 5 and so, I will come back to this issue later.

It is also important to point out that, apart from the percentage of looking time, Candan et al. (2012) also measured switches of attention, taken to be a signal of uncertainty during comprehension. They found that Mandarin children did not switch attention frequently as might be expected if they could not parse the sentence. In fact, their performance did not differ from English children in that respect, that is, as English children, Mandarin children were also very certain about matching sentences with corresponding scenes, which shows that Candan et al.'s results are not consistent across measures.

My study shared with Candan et al. (2012) the finding that vocabulary size did not relate in any systematic way to comprehension. As was observed in Candan et al. (2012), degree of comprehension (captured by the difference between test and control scores) was not significantly related to vocabulary size. The same conclusion was also reached in He and Lidz (2017), who investigated English-learning infants' early understanding of the link between the grammatical category *verb* and the conceptual category *event*; with the MacArthur CDI data collected, they found that although the 18-month-olds had a larger vocabulary on average than the 14-month-olds, they performed equally well, which indicates that the infants' ability in using the verb-event link to learn novel verbs did not depend on their productive vocabulary/verb knowledge.

Finally, the absence of age effects at 17 months suggests that the knowledge of canonical word order has been established before infants can actually produce an utterance, even in a language with null arguments and no overt case marking. Earlier works by Nespor, Guasti, & Christophe (1996) already claim that infants are sensitive to the OV/VO word order in the basis of prosodic cues (e.g. the weak-strong or strong-weak prosodic prominence reflects the language's VO/OV profile) in the early months of life, which is referred to as phonological bootstrapping. If we assume this hypothesis, it still cannot explain the results obtained from the present study, namely what prevents infants from interpreting an ungrammatical structure given the fact that OV (with *ba*) structures is also available in the input. In addition, sensitivity to the OV/VO is different from establishing the target word order, and in many cases prosodic representation does not fully reflect syntactic representation (Gerken, Jusczyk, & Mandel, 1994), and it is unclear how syntactic organization, as for example, displacement, can be directly read off from the phonological representation. In this respect, it would be interesting to test the comprehension of non-canonical word orders at roughly the same age and this was one of the motivations for the second experiment, described in the following chapter.

5 Early comprehension of Mandarin non-canonical word order

In this chapter, I report an empirical study using the preferential looking paradigm examining the comprehension of non-canonical word order including the morphosyntactic marker *ba* in Mandarin-speaking infants aged from 1;1 to 1;9.

5.1 Comprehension of non-canonical word order

There are few studies on the comprehension of non-canonical word order conducted on very young children. In a preferential looking task, Seidl, Hollich, & Jusczyk (2003) tested children exposed to English aged 15 and 20 months on the comprehension of *wh*-extraction from subject as in (67a) and object position as in (67b). Infants saw videos of two objects in interaction, such as a boy eating a pie. Then they saw both objects separately (boy and pie, for example) on different sides of the screen while hearing a subject (67a), object (67b), or locative where *wh*-question (67c):

(67) a. Who_i t_i ate the pie?

b. What_i did the boy eat t_i?

c. Where is the pie?

The authors found that 15-month old infants were able to comprehend (67a), while they argued that (67b) remains unparseable until around 20 months of age. In contrast, the 20-month-olds looked at the correct picture for both subject and object questions. Gagliardi, Mease, and

Lidz (2016) note that there are few data points in Seidl et al.'s (2003) design since there were only two trials for each type. They improved the design by offering the infants six trials in which an event occurred twice. The authors found a different pattern: 15-month old infants can process *wh*-extraction from both subject and object position and showed no evidence of the subject-object asymmetry reported in Seidl et al. (2003). Recently, Perkins and Lidz (2020) based on Gagliardi et al. (2016) showed that the correct interpretation of *wh*-extraction from object position and object relatives both correlate with productive vocabulary size in 15-month-olds. All these studies involve real verbs.

Lassotta, Omaki, & Franck (2014) (see also Lassotta, 2021), using eye-tracking techniques, tested 28 French-learning children aged 18-24 months (mean age 22 months) exposed to both SVO (as in (68a)) and clitic left dislocation (CLLD) with pseudo-verbs (as in (68b)), with theta-role reversal distractors.

(68) a. La fille dase le garçon.
DET girl PSEUDO-V DET boy
'The girl is dasing the boy.'

b. Le garçon, la fille le dase.
DET boy$_i$ DET girl CL$_i$ PSEUDO-V
'The boy, the girl is dasing him.'

The results showed that the looking time to the AGENT-first video increased (+18.7%) when infants heard SVO sentences, while it decreased (-12.3%) in OVclS condition. The proportion of looking time toward the AGENT-first video improved in the last window (in 46-49s) compared to the baseline in SVO condition; while it decreased compared to the baseline in OVclS condition in the last window. Thus Lassotta et al. (2014), Lassotta (2021) provided the first direct evidence that French infants can assign the correct PATIENT-AGENT-verb interpretation to non-canonical CLLD sentences with new verbs.

In the context of Mandarin, Hsu (2018), using the forced choice pointing paradigm, assessed Mandarin-speaking 2-year-olds' comprehension of the SVO (69a) and S*ba*OV constructions with overt subject (69b) and *ba*OV construction with a null subject (69c), with pseudo-verbs.

(69) a. Xiao-gou fo le xiao-mao.
little-dog PSEUDO-V PERF little-cat
'The puppy V-ed the kitty.'

b. Xiao-gou ba xiao-mao fo le
little-dog BA little-cat PSEUDO-V PERF
'The puppy V-ed the kitty.'

c. Ba xiao-mao fo le
BA little-cat PSEUDO-V PERF
'Someone V-ed the kitty.'

The results show that two-year-olds pointed to target trials 68% of the time for the canonical SVO construction, 67% for the S*ba*OV construction, and 67% for the subjectless *ba* construction, thus there were no comprehension differences among construction types. However, one potential confound in Hsu (2018)'s work is that both SVO and S*ba*OV are AGENT-first, and therefore could be interpreted correctly with an AGENT-first strategy as well.

5.2 Research questions

This line of research concerns the comprehension of non-canonical word order using eye-tracking measures adopting the IPLP for both infants and adults and taking as point of reference the experiment of Lassotta et al. (2014) and Lassotta (2021). Note that in Lassotta et al. (2014), Lassotta (2021) the canonical SVO order was represented by only two items, while there were four with the non-canonical OVclS order and the author found the bias towards the video illustrating the PATIENT-first interpretation was already present when infants heard canonical sentences in the baseline window. Thus, in my experiment, both canonical and non-canonical sentences have the same number of items (three test sentences in each condition).

In particular, I contrasted infants' capacity to interpret grammatical SVO, S*ba*OV and OS*ba*OV conditions with pseudo-verbs. Each condition was paired with two videos: one represented the target causative event, the other depicted the same event involving the same characters with theta-role reversal. This ensures that the focus is on the characters' roles, that is, on word order, and not on the lexical content of the verb (Gertner et al., 2006).

The rationale behind this is that if infants' sentence comprehension is guided by adult-like abstract syntactic knowledge, it is expected for them to show a preference for the scene with the first NP being the AGENT in the canonical SVO and non-canonical S*ba*OV condition, whereas, in the non-canonical OS*ba*OV, a preference for the scene with the first NP being the PATIENT is expected. Besides, if abstract syntactic knowledge of movement is absent, they will fail to parse the non-canonical S*ba*OV and OS*ba*OV condition. By contrast, if the infants' sentence comprehension is guided by heuristics like assigning the AGENT role to the first noun, then a preference for the scene with the first NP being interpreted as the AGENT is predicted in OS*ba*OV, SVO and S*ba*OV conditions. With respect to the lexical approaches, infants are predicted to perform at random, since all the sentences contain pseudo-verbs. Finally, since both S*ba*OV and OS*ba*OV are not frequent word orders in Mandarin, we can used them to test frequency-based models.

5.3 Infant Mandarin

This work is one of the few that deals with the very early acquisition of the non-canonical *ba* construction. From the corpus study in Chapter 3 we know that Mandarin-speaking children produce the *ba* construction at least when they approach their second birthday (at 1;8) and demonstrate productive knowledge at age 3-4 (see also Hsu, 2014). In this study I examine, for the first time, Mandarin-learning infants' comprehension of the *ba* construction using eye-tracking techniques.

5.3.1 Method

5.3.1.1 Participants

Twenty-four typically-developing Mandarin infants with a mean age of 17.5 months (*SD* = 2.2, range from 13 to 21 months, 12 boys and 12 girls) were tested. Thirteen additional infants participated in the study but they were not included in the results because of large errors in calibration (n = 5) or because of the infants' lack of compliance with the task (n = 8).

The child participants were recruited in the Danan Community Health & Services Centre, China, and had no reported history of speech, hearing or language disorders. Infants' vocabulary was measured using the Mandarin version of CDI (Hao et al., 2008). which include infant checklist (used for infants from 12 to 16 months of age) and toddler checklist (used for children between 17 and 30 months). Following Hao et al. (2008), for the toddler list, parents were only asked to indicate whether their children had ever said a word, thus no comprehension scores are available.

As can be seen in table 21, in my study, infants from 13-16 months (younger group) achieved a mean score of production of 9 words (SD = 7.1, range from 1 to 22) and a mean score of comprehension of 42 words (SD = 27.7, range from 2 to 88). Infants from 17-21 months (older group) achieved a mean score of production of 67 words (SD =74.9, range from 6 to 295).

Table 21: Infants' vocabulary

Months	Comprehension (words)		Production (words)	
	Mean	Range	Mean	Range
13-16	42 ($SD = 27.7$)	2-88	9 ($SD = 7.1$)	1-22
17-21	-	-	67 ($SD = 74.9$)	6-295

5.3.1.2 Materials and design

I tested both canonical SVO (70a) and non-canonical sentences involving the *ba* construction (70b,70c). Noun phrases referred to highly frequent animals in Chinese story telling (a cow, a dog, a duck, a rabbit, a sheep, and a horse) and all are included in both CDI checklists.

(70) a. Xiao-tu-zi tuan le xiao-gou.
little-rabbit PSEUDO-V PERF little-dog
'The bunny V-ed the puppy.'

b. Xiao-tu-zi ba xiao-ya-zi tuan le.
little-rabbit BA little-duck PSEUDO-V PERF
'The bunny V-ed the duckling.'

c. Xiao-ya-zi, xiao-tu-zi ba ta tuan le.
the-little-duck$_i$ little-rabbit BA it$_i$ PSEUDO-V PERF

'The duckling, the bunny V-ed it.'

A total of nine target items (three for each condition) were constructed as listed in Appendix A.4. Each sentence was associated to a synchronized pair of videos depicting the puppet characters carrying out an action that is not lexicalized in Mandarin (*tuān*, which means 'to put a colander on someone's head').

The test sentences were produced by a female native speaker of Mandarin (30yr). She was asked to produce all the sentences with target intonation. The recording was conducted in a sound-attenuated recording studio, recorded and digitized at 44.1 KHz, mono-channel. I used Adobe Premier Pro CC 2017 (v11.0.2) to edit the visual stimuli to make sure the two videos are synchronized, and Praat (Boersma & Weenink, 2005) was used to construct spoken utterances.

Infants were presented two synchronized videos, one video represented the target transitive event, the other depicted the same event involving the same characters with theta-role reversal. An example of test materials is illustrated in figure 20.

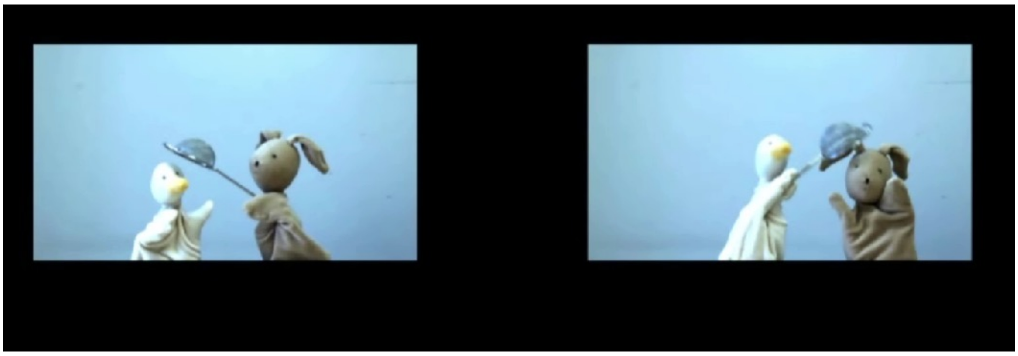


Figure 20: Test material, experiment 2

The nine test items were arranged in random order with the presentation of the target and reverse actions counterbalanced across the left and right sides of the screen.

5.3.1.3 Procedure

Participants' eye movements were recorded using a Tobii Pro X3-120 (with a sampling rate of 120 Hz) interfaced with a portable computer. Tobii Studio^{TM} (v3.4.8) was used as the platform for the recording and analysis of eye gaze data. Spoken utterances were presented to the participants through the two external speakers connected to the computer. Each child sat on his or her caregiver's lap during the whole length of the experiment. The distance between the participants' eyes and the monitor was about 60 cm. The caregivers were asked to close their eyes and listen to music played through headphones during the test trials so as not to guide their children towards any of the videos.

The participants were tested individually in a bright small room (300-350 Lux, Temperature 18-25^{o}C) without windows. Standard calibration was conducted before the experiment started, where a researcher, blind to the stimuli, observed the participants and clicked the mouse. Participants were asked to fixate on a grid of five calibration points in random succession. After having arrived at a satisfactory level of validity, the recording of eye movements took place from the onset of the training session and throughout. All the participants were simply told to listen to the spoken utterances while viewing the video.

As in the previous experiment on canonical word order, during the training session, first the participants went through a character-introduction phase (see Fig. 21). In the second phase, the participants were introduced to the character-identification phase (see Fig. 22).

Figure 21: The character-introduction video clip, experiment 2.

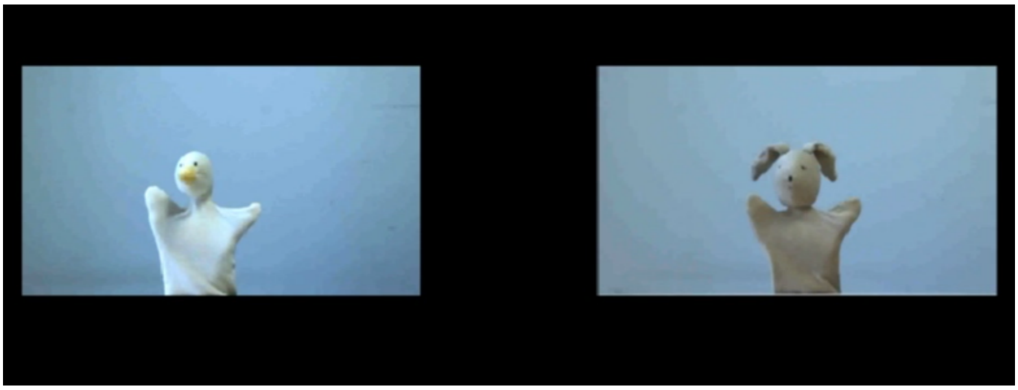


Figure 22: The character-identification video clip, experiment 2.

The third phase familiarized the participants with the novel action. As in the previous experiment, the novel action was presented once on the left (Fig. 23), once on the right of the screen (Fig. 24) and it was presented only in neutral frames such as *Bao-bao kuai kan, fa-sheng le shen-me?* 'Look, what happened?' so that later understanding of the test sentences cannot be attributed to learning during the training phase.

Figure 23: Novel action, left.

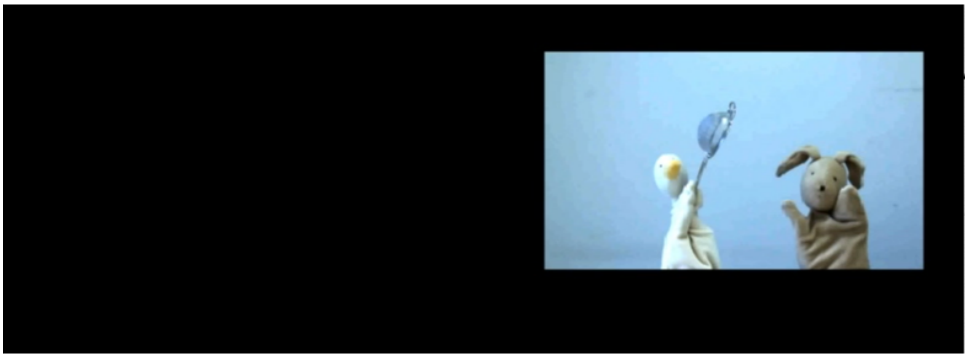


Figure 24: Novel action, right.

After the training session and a short transition cartoon with a song, the experimental session started (see Fig. 20). In the experimental session, all videos (9 pairs) started with an attentional sentence (e.g. *Bao-bao kuai kan, fa-sheng le shen-me?* 'Look, what happened?') as baseline, and then the experimental sentences were played three times. The recording of gazing time took place in four windows: the baseline and three consecutive exposures to the target sentence starting at 5, 10, 15 seconds. A blank screen (2s) appeared between experimental items, and after item 4, 6, 7 and item 8 a clip of the Teletubbies landscape was shown to keep the child's attention. The whole session lasted about 20 minutes. After the test session, the experimenter asked the infants' caregivers to fill out the Chinese version of CDI (Hao et al., 2008).

5.3.1.4 Predictions

If infants' sentence comprehension is guided by adult-like abstract syntactic knowledge of movement, I expect infants to show a preference for the scene with the 1st NP being the AGENT in the canonical SVO and non-canonical S*ba*OV condition, whereas, in the non-canonical OS*ba*OV, a preference for the scene with the 1st NP being the PATIENT is expected.

By contrast, if the AGENT-first conceptual bias guides sentence comprehension, a preference for the scene with the 1st NP being interpreted as the AGENT is predicted in SVO, S*ba*OV and OS*ba*OV conditions.

The item-based approach would predict that infants showed random behavior in all the three conditions, as the test sentences contain pseudo-verbs.

Frequency-based accounts would predict worse performance within the infrequent S*ba*OV and OS*ba*OV conditions compared to the frequent SVO condition.

5.3.1.5 Data treatment and analysis

Participants were excluded if there was more than 45% track loss from the onset; as a result, eight infants were excluded. Besides, five more infants were excluded because I were unable to calibrate them on the eye tracker, as already specified.

In analysing the eye movement data, I first categorically partitioned the data from the onset into four temporal RoIs: the baseline window (BS) and three consecutive exposures to the target sentence (First, Second, Third presentation of test sentences S1, S2, S3), starting at 5, 10, 15 seconds respectively. Then participants' fixations were coded in two scenes containing the target and non-target (i.e. reverse) events.

Mean looking times and proportion of fixations on each RoI in a specific temporal window were calculated. To provide an overview of the eye movement data, statistical analysis involved bivaried paired Student t tests for means' comparisons and LMMs for analysis of fixation proportions (see Zhan, 2018 for some discussion on this topic) using the lme4 (v1.1-12) package (Bates et al., 2013) from R (v3.2.5, R Development Core Team, 2017), and I used the Wald test to compute p-values for each fixed effect and the interactions.

Since fixation proportions follow a multinomial distribution rather than a normal distribution, they should be transformed to unbounded variables with empirical logit formula when using linear models (Barr, 2008). Following Barr (2008), fixation proportions were transformed using the formula in (71), where y is the number of fixations on each scene during a particular RoI, and n is the total number of fixations in that RoI.

(71) $$\eta = \ln\left(\frac{y + .5}{n - y + .5}\right)$$

I fit the data for each condition (SVO, S*ba*OV and OS*ba*OV) separately, focusing on the gaze duration toward the two critical scenes (target and reverse) in each condition. I accepted the simplified model if this model could explain the same variance as the full model. The model treated Scene (Target vs. Reverse) and RoIs (BS, S1, S2 and S3) as fixed effects, with random intercept and slope for participants and items.

Finally, I explored the effect of age, vocabulary and its interaction on proportion of looking time to the target scene using generalised linear models.

5.3.2 Results

The results of mean looking times of the canonical SVO structure appear in Table 22. They show that in the AGENT-first SVO condition, infants looked significantly longer at the target video than at the reverse video during the second (t(23) = 2.40, p = .025, Cohens d = .60) and the third exposure to the test sentence (t(23) = 3.99, p = .001, Cohens d = .66), reflecting their target SVO interpretation rather than OVS. No significant difference of looking time to the two videos was found in the baseline window, nor during the first presentation of the sentence.

Table 22: Mean looking times (in ms, *SD* in parentheses) across the four critical RoIs in the SVO condition, infants.
*p <.05, *** p <.001 (in bold)

SVO	Target	Reverse
Baseline (0-4s)	1299 (589)	1410 (820)
Sentence 1 (5-9s)	1511 (916)	1451 (853)
Sentence 2 (10-14s)	**1660 (837)***	**1139 (890)***
Sentence 3 (15-19s)	**1523 (865)*****	**1026 (628)*****

In the non-canonical S*ba*OV condition (see Table 23), infants still preferred looking at the target scene with the first NP being the AGENT during the second presentation of the sentence (t(23) = 2.19, p = .039, Cohens d = .51).

Table 23: Mean looking times across the four critical RoIs in the S*ba*OV condition, infants.
**p* <.05 (in bold)

S*ba*OV	Target	Reverse
Baseline (0-4s)	1931 (852)	1659 (746)
Sentence 1 (5-9s)	2273 (1103)	1829 (957)
Sentence 2 (10-14s)	**1944 (1193)***	**1396 (933)***
Sentence 3 (15-19s)	1595 (1079)	1900 (1204)

However, in the condition of objects in the topicalized position with the *ba* construction (see Table 24), infants disfavored the scene with the first NP being the AGENT and preferred looking at the opposite scene instead, i.e. the scene with the first NP being the PATIENT from the first presentation of the sentence (t(23) = 3.35, p = .003, Cohens d = .65), and the effect lasted until the end of the second exposure to the test sentence (t(23) = 2.08, p = .049, Cohens d = .57).

Table 24: Mean looking times across the four critical RoIs in the OS*ba*OV condition, infants.
**p* <.05, ** *p* <.01 (in bold)

OS*ba*OV	Target	Reverse
Baseline (0-4s)	1394 (665)	1163 (715)
Sentence 1 (5-9s)	**1867 (986)****	**1289 (785)****
Sentence 2 (10-14s)	**1888 (1102)***	**1323 (868)***
Sentence 3 (15-19s)	1364 (1060)	1290 (825)

The proportion of looking time to the target video (calculated over total looking time to the target and non-target videos) in the four critical RoIs is graphically represented in Figure 25.

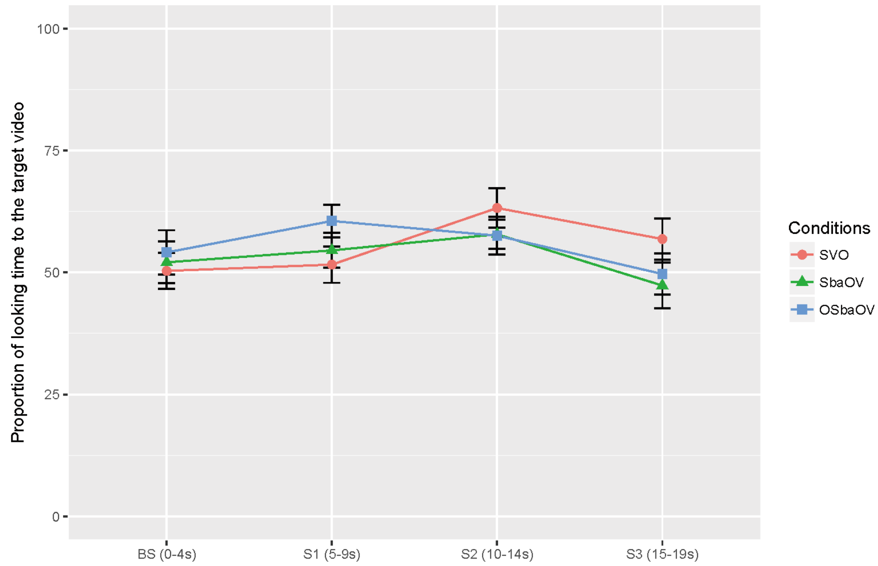


Figure 25: Proportion of looking time to the target video during the four critical RoIs in the three conditions, infants.

To assess the fixation patterns, LMMs were computed, and proportions were transformed with the empirical logit formula given in (71). Table 25 summarises the results of the model for the infant data in the SVO condition.

Table 25: Fixed effects from the best-fitting model of probability of looks to the target and reverse scenes in the SVO condition (empirical logit transformed), infants.
$*p < .05$, $**p < .01$, $***p < .001$ (in bold)

Fixed effects	Estimate	*SE*	*t* value
(Intercept)	−0.022	0.05	−0.238
Scene (Reverse)	0.045	0.150	0.303
S1	0.050	0.139	0.362
S2	0.361	0.139	**2.596***
S3	0.285	0.139	**2.050***
Scene (Reverse): RoIs (S1)	−0.100	0.191	−0.527
Scene (Reverse): RoIs (S2)	−0.722	0.191	−3.780 ∗ ∗∗
Scene (Reverse): RoIs (S3)	−0.570	0.191	−2.986 ∗ ∗

Formula in R: Proportion ∼ Scene*RoIs+(1+Scene+RoIs|Subject)+(1+Scene+RoIs|Item)

During the second and the third exposures to the target sentence, hearing SVO sentences caused the infants to look significantly longer at the target scene than at the reverse scene (β = .36, p = .01 for the second exposure and β = .29, p = .04 for the third exposure to the target sentence). The significant negative coefficient for the interaction between reverse scene and second and third exposures to the target sentence indicates that the probability of looking toward the reverse scene (i.e the scene with the first NP as PATIENT) decreased over time after hearing the SVO sentences.

As shown in Table 26, during the second exposure, hearing S*ba*OV sentences caused the infants to look marginally significantly longer at the target scene than at the reverse scene (β = .16, p =. 048). The negative coefficient for the interaction indicates that the probability of looking at the reverse scene decreased over time when the infants heard S*ba*OV sentences during the second presentation.

Table 26: Fixed effects from the best-fitting model of probability of looks to the target and reverse scenes in the S*ba*OV condition (empirical logit transformed), infants.
*p <.05 (in bold)

Fixed effects	Estimate	*SE*	*t* value
(Intercept)	0.093	0.120	0.777
Scene (Reverse)	-0.186	0.189	-0.985
S1	0.068	0.147	0.460
S2	0.162	0.147	**1.102***
S3	-0.207	0.147	-1.043
Scene (Reverse): RoIs (S1)	-0.136	0.208	-0.651
Scene (Reverse): RoIs (S2)	-0.325	0.208	**-1.558***
Scene (Reverse): RoIs (S3)	-0.413	0.208	-1.084

Formula in R: Proportion ~ Scene*RoIs+(1+Scene+RoIs|Subject)+(1+Scene+RoIs|Item)

As shown in Table 27, in the OS*ba*OV condition, the positive coefficient for the significant main effect of the first and the second exposures to the test sentence reflects the fact that the infants looked significantly more to the target scene when they heard the OS*ba*OV constructions compared to the target scene in the baseline window. The negative coefficient for the interaction confirms that, for infants, the probability of looking at the reverse scene decreased over time after hearing the first and the second presentation of the test OS*ba*OV sentences.

Table 27: Fixed effects from the best-fitting model of probability of looks to the target and reverse scenes in the OS*ba*OV condition (empirical logit transformed), infants.
**p* <.05 (in bold)

Fixed effects	Estimate	*SE*	*t* value
(Intercept)	0.165	0.104	1.582
Scene (Reverse)	-0.231	0.197	-1.171
S1	0.473	0.199	**2.376***
S2	0.494	0.208	**2.500***
S3	-0.029	0.270	-0.141
Scene (Reverse): RoIs (S1)	-0.347	0.270	**-1.282***
Scene (Reverse): RoIs (S2)	-0.334	0.270	**-1.235***
Scene (Reverse): RoIs (S3)	0.156	0.270	0.579

Formula in R: Proportion ~ Scene*RoIs+(1+Scene+RoIs|Subject)+(1+Scene+RoIs|Item)

I also explored the main effect of sentence condition and fixation to the areas with AGENT-first (see table 28) and I found a significant interaction between them, namely, hearing the SVO sentences caused the infants to look significantly longer at the scene with the first NP as AGENT (β = .65, t = 4.60, p < .001), the same happened when they heard the S*ba*OV sentences (β = .55, t = 3.93, p < .001). However, infants exhibited the opposite eye movement pattern when they heard the OS*ba*OV sentences (β = -.19, t = -2.75, p = .006), that is, hearing OS*ba*OV triggered more fixations to the scene with the first NP as PATIENT (see Figure 26). The gazing patterns provide evidence that they were able to identify the thematic roles encoded in word orders rapidly and effectively.

Table 28: Probability of looks (empirical logit transformed), infants
* p <.01, *** p <.001 (in bold)

Fixed effects	Estimate	*SE*	*t* value
(Intercept)	-0.19	0.07	**-2.75****
Condition (S*ba*OV)	-0.28	0.10	**-2.78****
Condition (SVO)	-0.32	0.10	**-3.25****
Agent-first (Yes)	-0.39	0.10	**-3.89*****
Condition (S*ba*OV):Agent-first (Yes)	0.55	0.14	**3.93*****
Condition (SVO):Agent-first (Yes)	0.65	0.14	**4.60*****

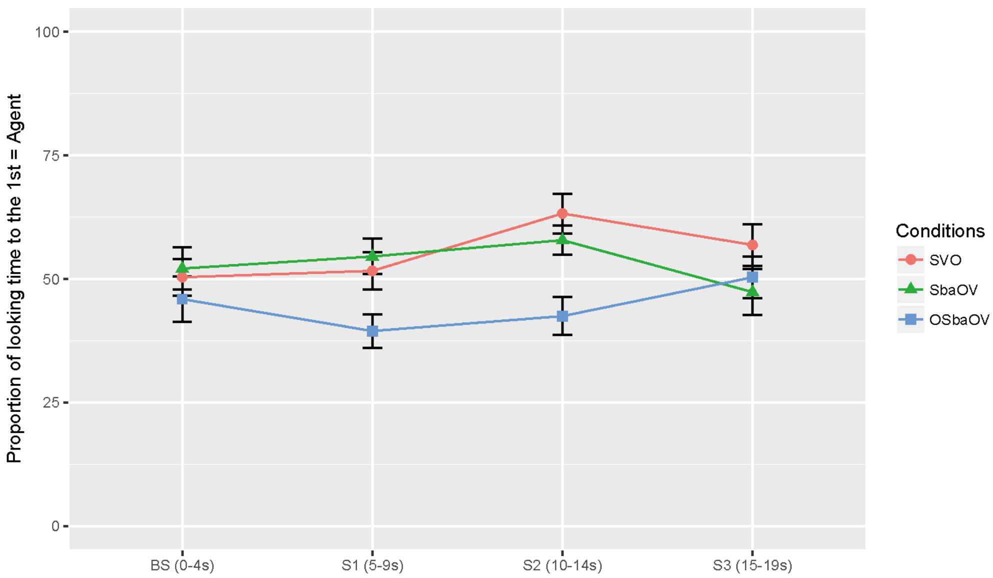


Figure 26: Proportion of looking time to the 1st NP = AGENT during the four critical RoIs in the three conditions, infants

Finally, generalized linear models with the proportion of looking time to the target scene as dependent variable were conducted to compare the main effect of age, vocabulary and their interaction. The results revealed that there was no main effect of age (β = .01, t = 1.08, p = .28) or vocabulary size (β = .003, t = .57, p = .56), and critically no interaction between vocabulary and condition (β = -.05, t = 1.11, p = .27), nor between age and condition (β = -.02, t = -.39, p = .70). Besides, the three-way interaction was not significant either (β = -.002, t = -.47, p = .64).

5.4 Adult Mandarin

The same experiment was conducted to discover adults' comprehension of non-canonical word orders with pseudo-verbs in Mandarin, while in many studies no results for adults are reported.

5.4.1 Method

5.4.1.1 Participants

Eighteen naïve Mandarin-speaking adults (mean age 29, *SD* = 7.7, range from 24 to 53 years old) participated in the experiment. They were recruited in Guiyang and Barcelona.

5.4.1.2 Materials, procedure and data treatment and analyses

I adopted the same material and analyses used for infants. For the adult participants, detected signal was more than 75% and no participants were excluded.

5.4.2 Results

For adults, the results in Table 29 and Table 30 show that, except for the baseline window, adults looked significantly longer at the target video than at the reverse video in the canonical SVO condition and non-canonical S*ba*OV condition during the first, the second and the third presentation of the sentence (all p's $< .001$).

Table 29: Mean looking times (in ms, *SD* in parentheses) across the four critical RoIs in the SVO condition, adults.
*** p <.001 (in bold)

SVO	Target	Reverse
Baseline (0-4s)	2225 (1342)	2941 (1544)
Sentence 1 (5-9s)	**4430 (1479)*** **	**2095 (1159)*** **
Sentence 2 (10-14s)	**5166 (1685)*** **	**1072 (1376)*** **
Sentence 3 (15-19s)	**5382 (1761)*** **	**1154 (1503)*** **

Table 30: Mean looking times (in ms, *SD* in parentheses) across the four critical RoIs in the S*ba*OV condition, adults.
*** p <.001 (in bold)

S*ba*OV	Target	Reverse
Baseline (0-4s)	2899 (1500)	2977 (1368)
Sentence 1 (5-9s)	**4304 (1472)*** **	**2119 (1048)*** **
Sentence 2 (10-14s)	**6126 (1793)*** **	**1050 (1113)*** **
Sentence 3 (15-19s)	**5964 (1819)*** **	**757 (845)*** **

In the OS*ba*OV condition (see Table 31), adults showed a latency during the first exposure to the test sentence, although they rapidly identified the target event during the second (t(17) = 3.09, p = .007, Cohens d = .73) and the third presentation of the sentence (t(17) = 2.70, p = .015, Cohens d = .64).

Table 31: Mean looking times (in ms, *SD* in parentheses) across the four critical RoIs in the OS*ba*OV condition, adults.
*** p <.001 (in bold)

OS*ba*OV	Target	Reverse
Baseline (0-4s)	2920 (1692)	2609 (975)
Sentence 1 (5-9s)	3206 (1560)	3121 (1443)
Sentence 2 (10-14s)	**4651 (1836)*** **	**2239 (1857)*** **
Sentence 3 (15-19s)	**4386 (1656)*** **	**2315 (1841)*** **

Proportion of looking time to the target video (calculated over total looking time to the target and non-target videos) in the four critical RoIs is graphically represented in Figure 27.

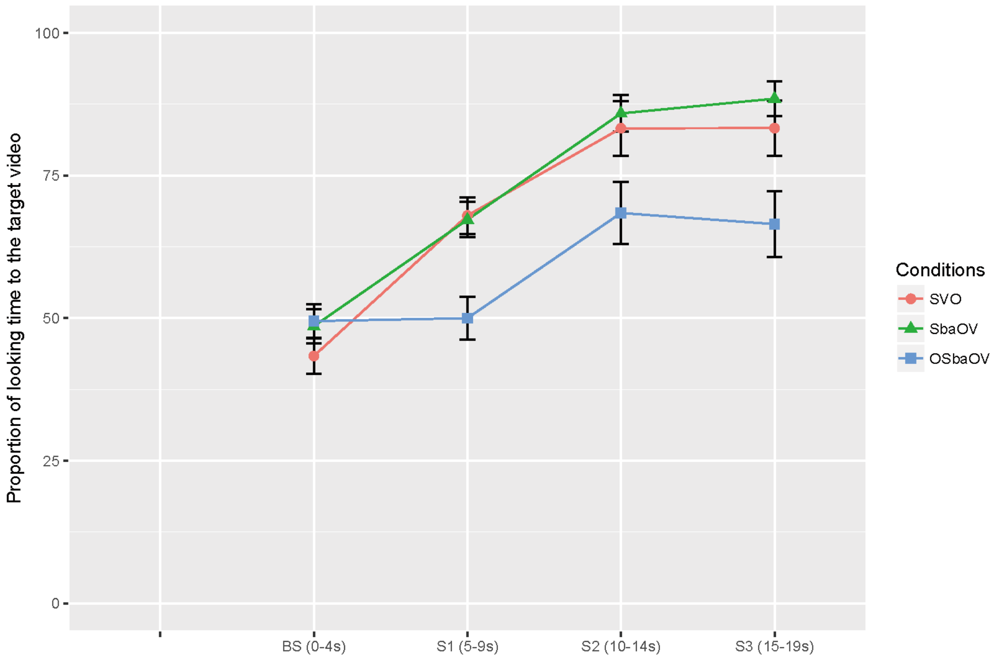


Figure 27: Proportion of looking time to the target video during the four critical RoIs in the three conditions, adults.

I assessed the gazing patterns of adults using the same methods as for infants. The negative coefficient for the interaction between the reverse scene and RoIs (see table 32 and table 33) indicates that, in both SVO and S*ba*OV conditions, the probability of looking at the reverse scene decreased after hearing the test sentences.

Table 32: Fixed effects from the best-fitting model of probability of looks to the target and reverse scenes in the SVO condition (empirical logit transformed), adults.
***p <.001 (in bold)

Fixed effects	Estimate	*SE*	*t* value
(Intercept)	-0.248	0.197	-1.263
Scene (Reverse)	0.415	0.349	1.191
S1	0.925	0.186	**4.964*** **
S2	1.851	0.186	**9.743*** **
S3	1.821	0.186	**9.774*** **
Scene (Reverse): RoIs (S1)	-1.769	0.262	**-6.753*** **
Scene (Reverse): RoIs (S2)	-3.549	0.262	**-13.553*** **
Scene (Reverse): RoIs (S3)	-3.577	0.264	**-13.527*** **

Formula in R: Proportion ∼ Scene*RoIs+(1+Scene+RoIs|Subject)+(1+Scene+RoIs|Item)

Table 33: Fixed effects from the best-fitting model of probability of looks to the target and reverse scenes in the S*ba*OV condition (empirical logit transformed), adults.
***p <.001 (in bold)

Fixed effects	Estimate	*SE*	*t* value
(Intercept)	-0.052	0.157	-0.333
Scene (Reverse)	0.104	0.277	0.376
S1	0.699	0.146	**4.804*** **
S2	1.696	0.146	**11.651*** **
S3	1.859	0.146	**12.769*** **
Scene (Reverse): RoIs (S1)	-1.398	0.206	**-6.793*** **
Scene (Reverse): RoIs (S2)	-3.392	0.206	**-16.478*** **
Scene (Reverse): RoIs (S3)	-3.717	0.206	**-18.058*** **

Formula in R: Proportion ∼ Scene*RoIs+(1+Scene+RoIs|Subject)+(1+Scene+RoIs|Item)

In the OS*ba*OV condition, there was no main effect of Scene when they heard the OS*ba*OV sentences for the first time (see table 34), which means that their interpretation was not significantly different from that of the baseline window as can be observed in Fig. 27, although, in the subsequent exposures, they interpreted the first noun as PATIENT efficiently.

Table 34: Fixed effects from the best-fitting model of probability of looks to the target and reverse scenes in the OS*ba*OV condition (empirical logit transformed), adults.
***p <.001 (in bold)

Fixed effects	Estimate	*SE*	*t* value
(Intercept)	-0.002	0.202	0.008
Scene (Reverse)	-0.003	0.339	-0.010
S1	-0.005	0.218	-0.023
S2	0.833	0.218	**3.819*** **
S3	0.768	0.218	**3.524*** **
Scene (Reverse): RoIs (S1)	0.010	0.308	0.032
Scene (Reverse): RoIs (S2)	-1.666	0.308	**-5.401*** **
Scene (Reverse): RoIs (S3)	-1.537	0.308	**-4.984*** **

Formula in R: Proportion ~ Scene*RoIs+(1+Scene+RoIs|Subject)+(1+Scene+RoIs|Item)

Surprisingly, adults showed a latency in the OS*ba*OV condition, which means that adults took longer than infants to look at the matching scene, although they identified the target event from the second presentation of the sentence. How could that be? This could be explained by the fact that the baseline sentence (e.g. *Fa-sheng le shen me* 'Look, what happened?'), a broad focus question, cannot be answered felicitously with a topicalized structure as illustrated in (72).

(72) a. Speaker A: Fa-sheng le shen-me?
happen PERF what
'What happened?'

b. Speaker B: #Xiao-ya-zi, xiao-tu-zi ba ta tuan le.
little-duck$_i$ little-rabbit BA it$_i$ PSEUDO-V PERF
'The duckling, the bunny V-ed it.'

This may cause confusion in adults due to pragmatical ill formedness. If this explanation is on the right track, by adding a more restrictive context, namely a narrow focus context (as in (73)), one would expect that adults would improve their performance. This was the motivation for a follow-up experiment.

(73) a. Speaker A: Xiao-ya-zi fa-sheng le
little-duck happen PERF
'What happened to the duckling?'
shen-me?
what

b. Speaker B: #Xiao-ya-zi, xiao-tu-zi ba ta tuan le.
little-duck$_i$ little-rabbit BA it$_i$ PSEUDO-V PERF
'The duckling, the bunny V-ed it.'

5.5 Follow-up Experiment

5.5.1 Method

Eleven Mandarin-speaking adults (mean age = 27, *SD* = 3.5, range from 22 to 32 years old) participated in this study and all were recruited in Barcelona. The test sentences contained both SVO in a broad-focus context (i.e. *Baobao ni kan, fa-sheng le shen-me?* 'Look, what happened?') and OS*ba*OV in a narrow-focus context (e.g. *Baobao ni kan, xiao-ya-zi fa-sheng le shen-me?* 'Look, what happened to the bunny?')

in the baseline sentence. The procedure was the same as in the previous experiment.

5.5.2 Results and discussion

The results (see Fig. 28) are consistent with my hypothesis. LMMs on total fixation time reveals that, in the OS*ba*OV condition, there was a significant effect of Scene (target vs. reverse) after the first (β = .98, t =3.03, p = .006), the second (β = 1.95, t =6.06, p < .001) and the third presentation of the sentence (β = 1.87, t =5.82, p < .001), with a significantly higher proportions of looks to the target PATIENT-first scene.

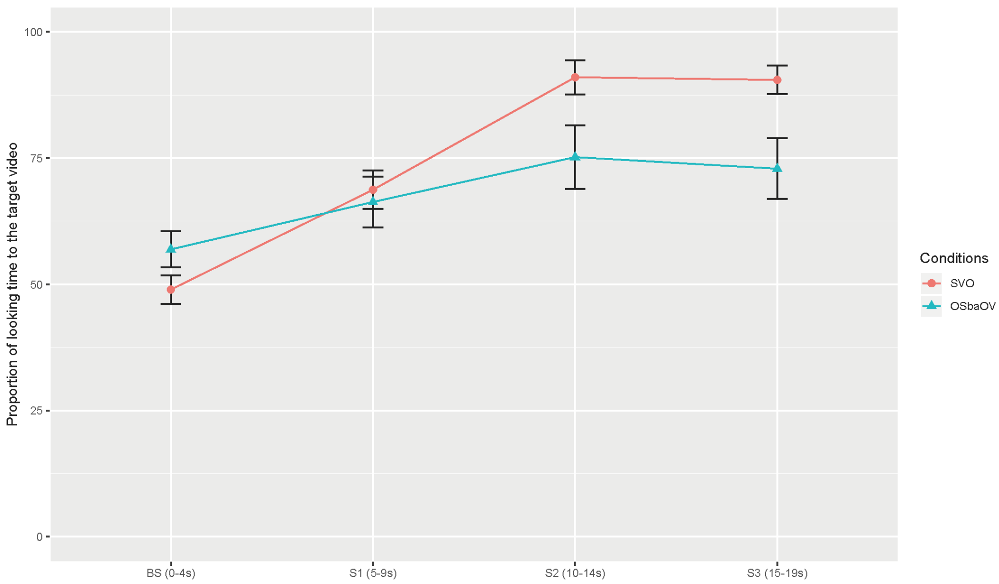


Figure 28: Proportion of looking time to the target video in the four critical RoIs in SVO with broad-focus context and OS*ba*OV with narrow-focus context, adults.

No latency was found in the first presentation of the sentence in contrast with the previous experiment with a neutral, broad-focus context as can be seen in Fig 29.

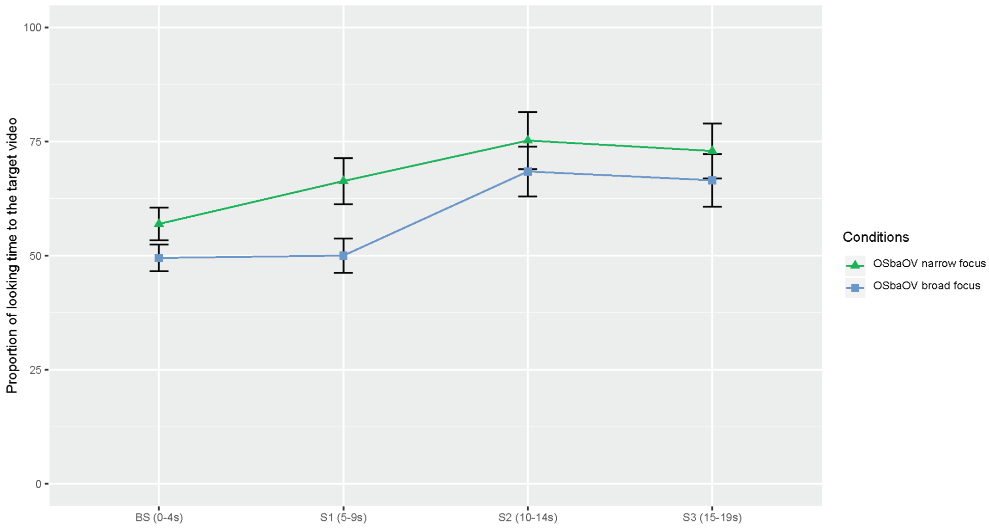


Figure 29: Proportion of looking time to the target video in the four critical RoIs in OS*ba*OV with narrow focus and broad focus context, adults.

5.6 Discussion

In this experiment I sought to investigate whether Mandarin-speaking infants are able to parse sentences with the object on the left periphery, instead of adhering to a putative universal AGENT-first parsing strategy. Using the intermodal preferential looking paradigm, I found that infants, as early as 17.5 months, can correctly assign the AGENT role to the first NP in both canonical SVO and non-canonical S*ba*OV constructions; in the OS*ba*OV condition, infants interpreted the first NP as PATIENT even during the first exposure to the sentence, reflecting their rapid fixation to the target interpretation. This shows that infants at this age are not simply sticking to an AGENT-first conceptual bias, as the PATIENT-first OS*ba*OV was parsed in a target-like manner, while an AGENT-first strategy would predict the theta-role universal interpretation. That is, even if infants begin with a preference for AGENT, by even before they actually utter a two-word sentence, they are able to override this preference and know at least some of the syntactic properties of the language they are exposed to.

An earlier study found that infants by 20 months of age are sensitive to both object (e.g. *what did the apple hit?*) and subject (e.g. *what hit the flower?*) *what*-questions, which implies movement of the object (Seidl et al., 2003). My findings extend the results to 17 months and suggest that, across languages, infant sentence comprehension is guided by adult-like abstract syntactic knowledge including movement from very early on.

In my study, higher vocabulary scores failed to predict the individual preferences for the target video, as would be expected if infants' interpretation of the sentences was dependent on their lexical knowledge. That is, 17-month-olds infants' ability in using the syntactic knowledge to learn novel verbs did not seem to depend on their productive vocabulary knowledge, in line with the results of the experiment in chapter 4 and He and Lidz (2017) as mentioned also in chapter 4, in which they found the vocabulary size does not predict the ability of linking pseudo-verbs to event concepts (as opposed to other concepts, e.g. objects)[19].

[19] Tardif (2015) by analyzing the naturalistic records of interactions between children and their caregivers at home found that Mandarin children start out with more verbs than common nouns compared to their English- (and Italian)-speaking peers, probably because subjects and objects, which are primarily nouns and pronouns, can be dropped in Mandarin. Thus there may be cross-linguistic differences between the vocabulary acquisition in different languages, which is a topic that goes beyond the scope of this dissertation.

The result obtained with the S*ba*OV construction is in sharp contrast with the result from the experiment in Chapter 4 which showed that children performed at random when they heard an ungrammatical SOV structure without *ba* (exemplified in (74)).

(74) *Xiao-niu shi-zi nui le.
little-cow lion PSEUDO-V PERF
'The calf the lion V-ed.'

As can be observed in Figure 30, first, infants exposed to Mandarin are sensitive to the presence of morphosyntactic marker *ba* from 17.5 months and they can use this knowledge to parse a sentence. Previous studies found that children as early as 25 months of age can use abstract grammatical features, like grammatical gender or aspect, to facilitate sentence comprehension and the sensibility to distinguish between content words and functional words could be present even earlier (see Dye et al., 2019 for a review). For example, French-learning toddlers can use gender (Van Heugten & Shi, 2009) and number features (Robertson, Shi, & Melançon, 2012) of determiners to identify the novel nouns. Santelmann and Jusczyk (1998) showed that 18-month-olds presented a significant listening preference when the auxiliary *be* was followed by a main verb ending with -*ing* (e.g. *Everybody is baking bread.*) rather than for sentences which contained an ungrammatical combination of the modal auxiliary *can* and a main verb ending with -*ing* (e.g. *Everybody can baking bread*). My study shows that infants are sensitive to an even less frequent functional head, *ba*.

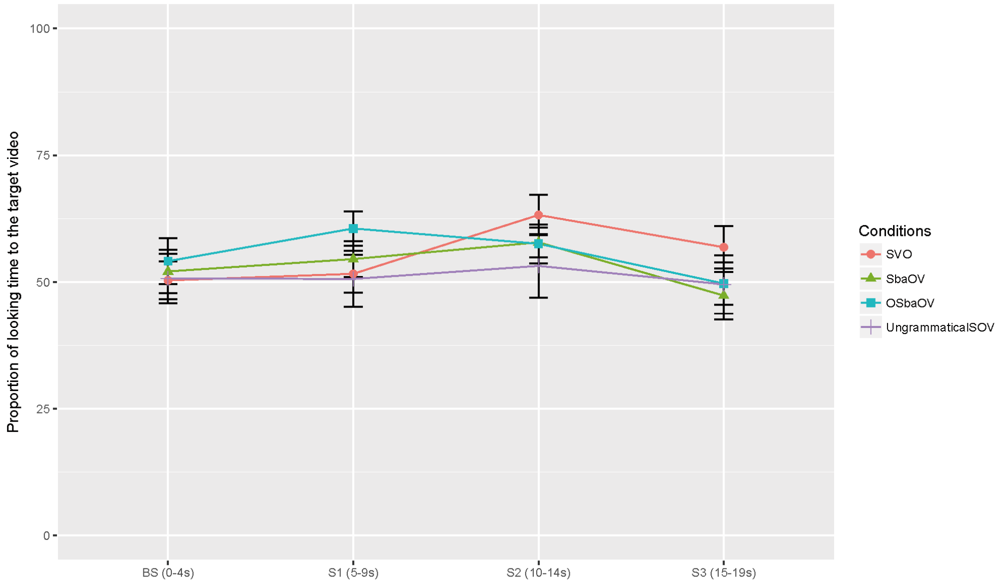


Figure 30: Proportion of looking time to the target video during the four critical RoIs in the four conditions, infants.

Second, in the experiment on comprehension of canonical word order, infants saw two simultaneously videos, one illustrated a transitive event and another described a reflexive event. The infants' correct interpretation of grammatical SVO sentences, but not ungrammatical SOV sentences in Mandarin with pseudo-verbs, lead me to conclude that infants at 17 months can use word order to infer the semantic meaning (i.e. syntactic bootstrapping). My results in the present experiment extend this claim: infants at 17 months are able not only to reconstruct something about the semantics given the relevant syntactic environment, but also they can assign the correct thematic roles to each NP, as in the present experiment infants were paired with two scenes with theta-role reversal, both depicting transitive events. If infants just relied on the semantic structures of the *ba* construction, which is associated with highly transitive events (Sybesma, 1992; Wang, 1945, 1987), they would show random behavior in the absence of syntactic knowledge.

Of course, one might argue that infants in my experiment can use only partial knowledge to interpret the test sentences involving moved objects without having knowledge of the movement operation. The experiment on canonical word order has already indicated that, by 17 months, infants have already shown knowledge of argument structure. Thus for a sentence like (70b) (repeated here as (75) for convenience), they could recognize a gap by noticing that after the verb there is an expected syntactic argument that is missing (e.g., *the bunny ba the duckling tuan-ed* ___) and this missing argument should be an object, and the heuristic parser would then search the discourse context for a referent that could fill out the gap, and normally the second NP of a canonical sentence corresponds to the object.

(75)	Xiao-tu-zi	ba	xiao-ya-zi	tuan	le.
	little-rabbit	BA	little-duck	PSEUDO-V	PERF

‘The bunny V-ed the duckling.’

Thus, this strategy would allow for the correct interpretation of (75) without forming the link between the filler and the gap and would not require children to parse the filler to arrive at the correct interpretation. In fact a previous study is argued to show that 15-month-olds rely on partial knowledge to correctly interpret both subject and object relative clauses (Gagliardi et al., 2016), while 20-month-olds had more difficulty comprehending object relative clauses like (76), and the authors attributed this phenomenon to the fact that 15-month-olds did not really understand the object relative clause, but used the heuristic strategy to parse the sentence.

(76) Show me the dog that the cat bumped.

If infants recognize *bump* is a transitive verb (during the training phase), they will notice that after *bump* there is an expected syntactic argument that fails to occur, then what they need to do is to search this missing argument through the discourse context without establishing the filler-gap link.

Nevertheless, we can rule out this possibility here. First, if 17-month-olds use partial knowledge rather than adultlike grammatical knowledge to interpret S*ba*OV sentences, then why did they fail to show any preferences when encountering ungrammatical SOV sentences in the first experiment? Second, if they use the heuristic strategy (i.e. the second NP is the object) to parse OS*ba*OV sentences, that would lead them to interpret the second NP as object. However, they correctly interpreted the second NP as subject in such condition.

Thus, it seems like infants at 17 months have already acquired grammatical knowledge of movement. Recall that ungrammatical SOV would also involve movement, but infants would never parse that, indicating the integration of UG and PLD. In addition, the absence of age effects seems to suggest, again, that the parameter(s) related to the movement of the object have been fixed earlier than 17 months, which is consistent with the claim that (internal and external) Merge is a crucial operation of the human language faculty, part of biological endowment (Berwick et al., 2013) and, furthermore, that it is active from early on, before 17 months. Just like *why*-questions in English do not include movement (Rizzi, 2001), and children acquire *why*-questions quite late (Thornton, 2008, 2016), child performance is not ruled by a simple complexity measure based on presence/absence of movement.

Moreover, note that the *ba* construction is scarce in the input (only

3.1% in child-directed speech, see Chapter 3) and an analysis of the CHILDES corpus of child-directed speech showed that OS*ba*-cliticV sentences are present in only 1% of sentences in the input (Yeh, 2015), which means that children get very little exposure to the construction. This is at odds with the predictions of frequency-based models (see also Hsu et al., 2019 as mentioned in chapter 4). The same happens in French clitic left dislocation (CLLD); according to Lassotta (2021), object fronting represents about 7.8% of all clauses in the French Lyon corpus (Demuth & Tremblay, 2008), and the object topicalization structure used in Lassotta et al. (2014) and Lassotta (2021) only occurs in 0.1% (27/25310 clauses).

Thus, despite the very low frequency and the additional complexities of the OS*ba*OV structure, where the object has been topicalized in the left periphery and is coindexed with a resumptive clitic pronoun in preverbal position (as exemplified in (70c), repeated here as (77) for convenience), children still identified the target event very quickly (see the heat map in Fig. 31), a result similar to that of Lassotta et al. (2014) for French CLLD.

(77)	Xiao-ya-zi,	xiao-tu-zi	ba	ta	tuan	le.
	the-little-duck$_i$	little-rabbit	BA	it$_i$	PSEUDO-V	PERF

‘The duckling, the bunny V-ed it.’

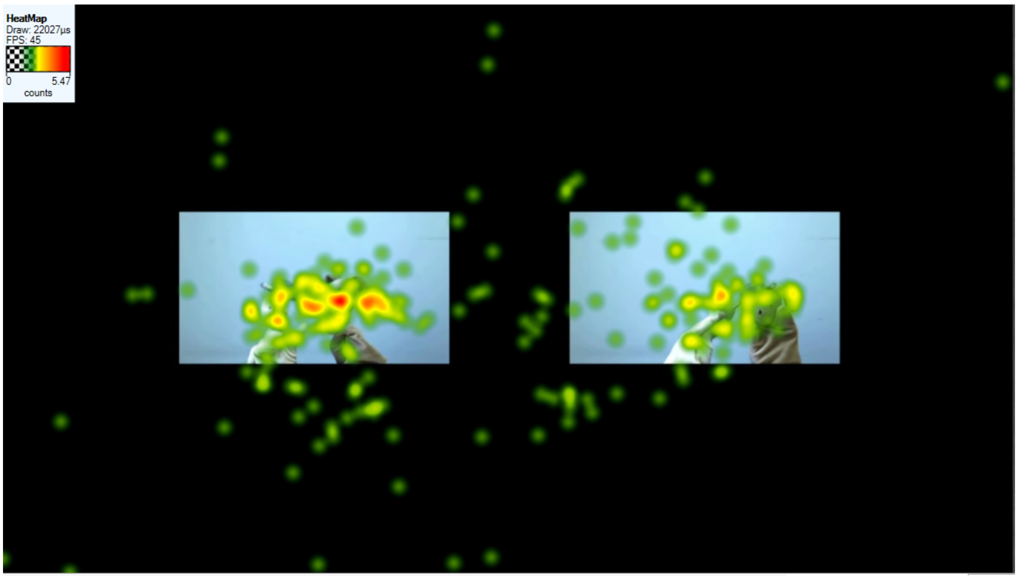


Figure 31: Heat map for the test sentence *Xiao-ya-zi, xiao-tu-zi ba ta tuan le* 'The duckling, the bunny V-ed it'. Red indicates the highest number of fixations or the longest time, and green the least. Left video represents the target scene and the right video represents the reverse scene.

The quick identification of the topicalized structure is consistent with the idea that there is no delay in A' movement in child grammar (Babyonyshev et al., 2001), and that children's constructions encode information structure (Belletti & Manetti, 2018; De Cat, 2009) at an early age.

My results seem to be inconsistent with the predictions made based on child Relativised Minimality (Friedmann, Belletti, & Rizzi, 2009), in the sense that the OS*ba*OV structure could give rise to minimality effects if analysed as (78), and yet it is well understood:

(78)	O	S	ba	O	V (Ø)
	[+Top, +NP]	[+NP]	BA	[+Top, +NP]	V

However, Friedmann et al. (2009) do not discard the idea that child Relativised Minimality is a processing effect instead of grammatical issue.

Now the question is why the processing limitation does not influence Mandarin-speaking infants in this study still needs to be explained, but this is beyond the scope of this dissertation.

Prosodic features may help identify the topic. It is generally agreed that canonical topic structures (like *xiao-ya-zi* 'duckling' in (77)) are outside the TP and pause is the most important and reliable indicator of topichood in Chinese (Chao, 1968). OS*ba*OV is the only construction in the experiment that requires a pause between the first NP (i.e. the topicalized object) and the second argument (i.e. the subject). Then infants may know that the NP before the pause is the topic. Moreover, pause *per se* may also help acquire the construction. In my study, the average duration of the pause in the OS*ba*OV construction between the first NP and the second NP was 373 ms. Recent work by Marchetto and Bonatti (2015) has found that when the auditory offset of one word and the onset of the next word were separated by 200 ms of silence, 12-month-old infant can build word-internal structure, while they fail to grasp this structure when the pauses marking boundaries is eliminated. Although this only applies to the lexical level, we cannot discard the possibility that a pause may favour the recognition of certain syntactic structures. While only speculations are possible with the data available at the moment, it would certainly be worth comparing performance with OS*ba*OV with and without a pause.

The fact that adults showed a latency in the OS*ba*OV condition in the broad focus context can be indeed explained by pragmatic considerations. In the OS*ba*OV condition, the topicalized referent (i.e. the object) must be somehow connected to the discourse context, and an out-of-the-blue, all-new context (i.e. the broad focus *what happened?*) does not licence a topic-comment structure. This is corroborated by the follow-up experiment providing a narrow focus context, as well as a study on Cantonese. Lai (2018) shows that, in an acceptability judgements task, native Cantonese speakers disfavoured the topicalization of the object (i.e. OSV constructions) in the broad-focus context like the one in my study (*Fa-sheng le shen me?* 'Look, what happened?'). Since Mandarin and Cantonese share variable word order for pragmatic functions, it is reasonable to consider that the Cantonese adult results are in line with those of Mandarin.

Combing the follow-up results with the original ones, it seems that infants are likely to ignore the infelicity of the answer, and rely mostly on the syntactic knowledge of the structure, which would show that perhaps children are pragmatically more tolerant than adults. If the conjecture is on the right track, then the difference between adults and children in the present experiment is not grammatical knowledge, but growing pragmatic sophistication.

6 Finding closure

6.1 The end: To be continued

In the course of the present work, a number of questions were raised that need to be addressed and call for future research.

In Mandarin, there is a small set of verbs which can have both transitive and intransitive interpretations, moreover, in those structures, the AGENT of the intransitive sentence should perform a different action from that of the transitive sentence. An example of such alternation can be found in (79).

(79) a. Xiao-ming cang qi-lai le.
Xiaoming hide DC PERF
'Xiaoming has hidden.'

b. Xiao-ming cang qi-lai le yi ben shu.
Xiaoming hide DC PERF one CLASS book
'Xiaoming has hidden a book.'

In (79a), Xiaoming himself has hidden, whereas in (79b), he did not hide. If null objects are allowed in child early grammar, I should expect that, given the right context, both interpretations like 'Xiaoming has hidden' and 'Xiaoming has hidden some previously mentioned object' should be available to them. So in a future study, I could conduct another IPLP task. In the test, if the child accepts (79a) in a context where Xiaoming has hidden some previously mentioned object while he himself did not hide, I can assume that her interpretation is based on the fact that Chinese allows null objects, as illustrated in (80):

(80)
- **Context**: Xiaoming has hidden the book in the closet. (The book is mentioned in the immediately preceding context).
- **Test sentence**: *Xiaoming cang qi-lai le* 'Xiaoming has hidden'.

  a) **If null objects are not allowed in child grammar**: she will look at the picture in which Xiaoming himself has hidden.

  b) **If null objects are allowed in child grammar**: she will look at the picture in which Xiaoming has hidden the book in the closet.

Besides, since NP-NP-V constructions often co-occur with passive marker *bei* as I mentioned in section 2.3.3, and earlier work has showed that participants relied overwhelmingly on *bei* in choosing the second NP as the AGENT (Li et al., 1993), a follow-up experiment on long passive O*bei*SV (like (81)) might be interesting to conduct in future work.

(81) Xiao-ya-zi bei xiao-tu-zi tuan le.
little-duck BEI little-rabbit PSEUDO-V PERF
'The duckling has been V-ed by the bunny.'

Although it is often stated that English-speaking children do not master passive constructions early (Borer & Wexler, 1987; Horgan, 1978), as the passive structure in adult grammar includes VP-smuggling or movement across a phase, which could be problematic for young children under various accounts (see grammatical accounts like Wexler's (2004)

Universal Phase Requirement or Snyder and Hyams' (2015) smuggling account), it seems that Chinese passive will present a different picture, as it is tangential to the issue of VP-smuggling or movement across a defective phase. Several studies have shown that Mandarin passives are similar to the English *get*-passives, rather than BE-passives (Huang et al., 2009). Previous studies have found that *get*-passives are common in children's production (Crain, Thornton, & Murasugi, 2009; Harris & Flora, 1982), and are also comprehended quite well on account of the fact that they do not involve the syntactic operations underlying *BE*-passives.

The child Relativised Minimality issue mentioned above is clearly a topic for future research since my results seem to be at odds with the predictions made based on it, in the sense that the OS*ba*OV structure would be predicted to give rise to minimality effects.

Another issue that emerges in relation to using eye-tracking in developmental cognitive neuroscience, the problems associated with collecting eye-tracking data and analyzing them with infants, has to be taken into account, since the accuracy and precision of eye-tracking data collected from infants are often lower than the accuracy and precision of eye-tracking data obtained from adults (see Hessels & Hooge, 2019 for discussion); it would be interesting to investigate word order combining with other measurement techniques, as EEG or fMRI in the very same experiments.

6.2 Conclusion

The results presented in this dissertation provide evidence that infants are sensitive to word orders of their target language and can use them in comprehension during very early developmental stages surrounding 17 months (or even before). My first experiment concludes that infants acquiring Mandarin preferentially look at transitive scenes when they hear grammatical NP-V-NP frames, whereas no significant preference is observed when infants are confronted to ungrammatical NP-NP-V frames. My second experiment leads us to conclude that infants not only are aware that their target language allows flexibility in word order, but also are sensitive to certain functional heads of their target language (as *ba* in my study) based on the PLD.

Now, let us go back to my research questions: do infants have abstract grammatical knowledge of word order? If they do, at which age? In the two eye-tracking experiments reported, no age effect has been found within the age range tested: comprehension for the targeted age range does not relate to age, not does it relate to vocabulary size. The absence of the age effect seems to suggest that both canonical word order and non-canonical word orders have been acquired earlier than 17 months, independently from input frequency. In order to be able to comprehend utterances that they have never heard before, such as those in my experiments, infants need to have abstract syntactic representations. Besides, the corpus study shows that at 1;8 or MLU stage II (1.75-2.25), children show no target-deviant productions with respect to the production of null and overt subjects and objects.

These findings are inconsistent with the lexical approaches and with the variational model if they apply at age 17 months, since infants can parse grammatical SVO, S*ba*OV and OS*ba*OV constructions with pseudo-verbs at 17 months. Moreover, there is no variational stage where null and overt objects are seen to be competing, since all the children that I observed produced overt objects and null objects like adults from the first occurrence. This does not mean that the acquisition of the lexicon plays no role in early language acquisition. Of course, lexical learning does play a role in sentence comprehension; as stated by Gleitman, Liberman, McLemore, & Partee (2019), it is hardly imaginable that one can start syntactic bootstrapping if one does not even know where the subject of the sentence is. That is, to trigger syntactic bootstrapping, one needs to learn a few nouns as an anchor, which makes syntactic bootstrapping possible.

Infants can parse sentences involving grammatical movement (S*ba*OV and OS*ba*OV), but they failed to do so when the movement is illegitimate (as SOV). In addition, they can use adult-like syntactic knowledge to find the unpronounced gap in S*ba*OV sentences while, in contrast, L2 adults exhibit a huge problem with this kind of constructions (Huang & Yang, 2004; Wen, 2012), and native-like performance is only found at the advanced level (Wen, 2012). Besides, those preference patterns were observed with pseudo-verbs, of which the infants had no previous knowledge. Infants' performance on PATIENT-first OS*ba*OV construction, involving object movement, shows that they do not make use of a universal bias like assigning the AGENT role to the first NP or resort to some partial knowledge, rather, from age 1;5 at the latest, they have abstract knowledge of their target language, and they can use this knowledge to comprehend sentences. Their response pattern thus appears to be grammar-based.

These findings are consistent with the evidence already gathered on

Indo-European languages (French, Hindi-Urdu), albeit for a slightly older age (19 months, rather than the 17 months of the participants in the present study). What is also new here is that they come from a language which displays word order variation and the presence of null arguments as well as the lack of agreement and case marking.

The fact that infants at 17 months are sensitive to the different word orders seems to indicate that infants are not mis-setting parameters related to it. The results are consistent with Gibson and Wexler (1994). The learner identifies a default grammar, and only needs to change the value of the default grammar if the later does not allow her to parse the current sentences in the input. It would be consistent with the variational model if that model characterized an early stage (i.e. before 17 months of age), but that is inconsistent with Yang's claim (2002; 2004) that a variational stage is visible past the two-word stage.

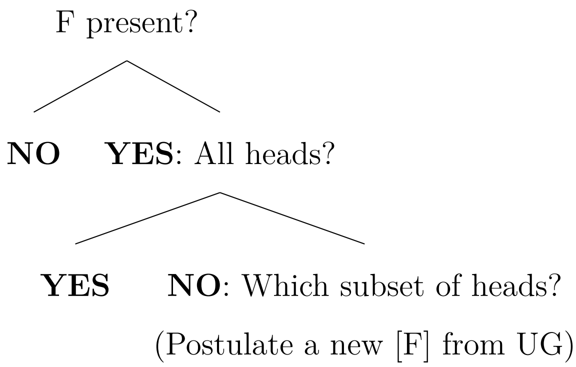
In addition, word order invokes the notion of 'macroparameter', which characterizes languages into 'types' (Baker, 2001; Richards, 2008; Roberts & Holmberg, 2010) and is set early (Wexler, 1998). Biberauer (2017) highlights that certain linguistic properties, such as systematic silence (null arguments and ellipsis) and movement may alert the acquirer to pay more attention to functional features. Consider the sentence S*ba*OV and OS*ba*OV in (75) and (77), repeated below.

(82) a. Xiao-tu-zi ba xiao-ya-zi tuan le.
little-rabbit BA little-duck PSEUDO-V PERF
'The bunny V-ed the duckling.'

b. Xiao-ya-zi, xiao-tu-zi ba ta tuan le.
the-little-duck$_i$ little-rabbit BA it$_i$ PSEUDO-V PERF
'The duckling, the bunny V-ed it.'

The contrast between the unmarked SVO structures and these object movement structures requires reference to the formal feature [+ Topic] and *ba*, thus driving the acquirer to pay more attention to them. In addition, according to the learning path proposed by Biberauer (2019) (see also Maximise Minimal Means), for economic reasons, the child by default will postulate as few formal features as possible from UG to account for the input, but once the child proposes a formal feature [F], s/he will maximise already-postulated features to all heads. In this sense, the acquisition follows a general learning path NO > ALL > SOME pattern as in (83) from Biberauer (2019, p. 60) (see also Biberauer & Roberts, 2016).

(83)



By postulating a path like (83), the acquisition order becomes relevant: macroparameters will be set first. In this dissertation, I have shown that SVO, S*ba*OV and OS*ba*OV orders are available at age 17 months, meaning that the formal feature triggering the *ba* construction has been acquired well before that age.

Throughout the second semester of life babies polish their sensitivity to particular phonological features of their native language (Kuhl, Ramírez, Bosseler, Lin, & Imada, 2014; Pons, Lewkowicz, Soto-Faraco, & Sebastian-Gallés, 2009). In addition, they may be able to recognize their mother's voice as far back as uterine life (Jardri et al., 2012; Kisilevsky et al., 2009) and newborns' cry melody even mirrors their ambient language, an ability that apparently acquired *in utero* (Mampe, Friederici, Christophe, & Wermke, 2009). Then there is no reason not to hypothesize that starting from an early age or maybe in parallel to the stage when they show sensitivity to native phonotactics, babies also become sensitive to some macro-structures or features of their native language, such as word order. As pointed out by Golinkoff et al. (2013), if so much is happening prior to the production of the first sentence, current theories of language acquisition need to change to accommodate this precocity.

References

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2001). What pre-school children do and do not do with ungrammatical word orders. *Cognitive Development*, 16(2), 679–692.

Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). Graded representations in the acquisition of English and German transitive constructions. *Cognitive Development*, 23(1), 48–66.

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematizartion in a usage-based account of syntactic acquisition. *The Linguistic Review*, 23, 275–290.

Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language*, 26(2), 339–356.

Ambridge, B., Bidgood, A., Pine, J. M., Rowland, C. F., & Freudenthal, D. (2016). Is passive syntax semantically constrained? Evidence from adult grammaticality judgment and comprehension studies. *Cognitive Science*, 40(6), 1435–1459.

Ambridge, B., & Lieven, E. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge: Cambridge University Press.

Ambridge, B., & Lieven, E. (2015). A constructive account of child language acquisition. In B. MacWhinney & W. 'Grady (Eds.), *The handbook of language emergence* (p. 478-510). New York: Wiley.

Aslin, R. (2012). Infants eyes: A window on cognitive development. *Infancy*, 17(1), 126–140.

Aslin, R., & McMurray, B. (2004). Automated corneal-reflection eye tracking in infancy: Methodological developments and applications to cognition. *Infancy*, 6(2), 155–163.

Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.

Babyonyshev, M., Ganger, J., Pesetsky, D., & Wexler, K. (2001). The maturation of grammatical principles: Evidence from Russian unaccusatives. *Linguistic Inquiry*, 32, 1–44.

Baker, M. (1988). *Incorporation.* Chicago: Chicago University Press.

Baker, M. (2001). *The atoms of language.* New York: Basic Books.

Baker, M. (2008). *The syntax of agreement and concord.* New York: Cambridge University Press.

Banks, M., Aslin, R., & Letson, R. (1975). Sensitive period for the development of human binocular vision. *Science*, 190(4215), 675-677.

Bannard, C., Lieven, E., & Tomasello, M. (2009). Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences*, 106(41), 17284–17289.

Barbosa, P. (2019). *pro* as a minimal nP: Toward a unified approach to pro-drop. *Linguistic Inquiry*, 50(3), 487–526.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59, 457-474.

Bates, D., Maechler, M., & Bolker, B. (2013). *lme4: Linear mixed-effects models using s4.* R package version 3.2.5.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using eigen and s4.* R package version 3.5.2.

Bates, E., Elman, J., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In L. Gleitman & E. Wanner (Eds.), *Language acquisition: The state of the art* (p. 173-218). Cambridge: Cambridge Univesity Press.

Belletti, A., & Manetti, C. (2018). Topics and passives in Italian-speaking children and adults. *Language Acquisition*, 26(2), 153–182. (DOI: https://doi.org/10.1080/10489223.2018.1508465)

Bencini, G., & Valian, V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, 59, 97–113.

Berwick, R., Beckers, G., Okanoya, K., & Bolhuis, J. (2012). A bird's eye view of human language evolution. *Frontiers in Molecular Neuroscience*, 4:5, 1-25. (DOI: https://doi.org/10.3389/fnevo.2012.00005)

Berwick, R., & Chomsky, N. (2011). Biolinguistics: The current state of its evolution and development. In A. di Sciullo & C. Boeckx (Eds.), *Biolinguistic investigations* (p. 19-41). Oxford: Oxford University Press.

Berwick, R., & Chomsky, N. (2017). *Why only us. Language and evolution*. Cambridge, MA: MIT Press.

Berwick, R., Chomsky, N., & Piattelli-Palmarini, M. (2013). Poverty of the stimulus stands: Why recent challenges fail. In M. Piattelli-Palmarini & R. Berwick (Eds.), *Rich languages from poor inputs* (p. 19-42). New York: Oxford University Press.

Bever, T. (1970). The cognitive bias for linguistic structures. In J. Hayes (Ed.), *Cognition and the development of language* (p. 279-362). New York: Wiley.

Biberauer, T. (2017). Particles and Final-over-Final constraint. In M. Sheehan, T. Biberauer, A. Holmberg, & I. Roberts (Eds.), *The Final-over-Final Condition* (p. 187-296). Cambridge, MA: MIT Press.

Biberauer, T. (2019). Factor 2 and 3: Towards a principled approach. *Catalan Journal of Linguistics*, Special Issue, 45–88. (DOI: https://doi.org/10.5565/rev/catjl.219)

Biberauer, T., Holmberg, A., & Roberts, I. (2007). Disharmonic word-order systems and the Final-over-Final-Constraint (FOFC). In A. Bisetto & F. Barbieri (Eds.), *The Proceedings of XXXIII Incontro di Grammatica Generativa* (p. 86-105). Bologna: Università di Bologna.

Biberauer, T., Holmberg, A., & Roberts, I. (2008). Structure and linearization in disharmonic word orders. In C. Chang & H. Haynie (Eds.), *The proceedings of the 26th western coast conference on formal linguistics* (p. 96-104). Somerville, MA: Cascadilla Proceedings Project.

Biberauer, T., Holmberg, A., & Roberts, I. (2014). A syntactic universal and its consequences. *Linguistic Inquiry*, 45(2), 169–225. (DOI: https://doi.org/10.1162/LING_a_00153)

Biberauer, T., & Roberts, I. (2016). *Emergent parameters and syntactic complexity: New perspectives.* Paper presented at the 49th Annual Meeting of the Societas Linguistica Europaea, Università degli Studi di Napoli Federico II, Naples.

Boersma, P., & Weenink, D. (2005). *Praat: Doing phonetics by computer.* http://www.praat.org/. (Accessed: 06/06/2020)

Bolker, B., Brooks, M., Clark, C., Geange, S., Poulsen, J., Stevens, M., & White, J.-S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24(3), 127-135. (DOI: https://doi.org/10.1016/j.tree.2008.10.008)

Borer, H. (1984). *Parametric syntax.* Dordrecht: Foris.

Borer, H., & Wexler, K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.), *Parameter setting* (p. 123-172). Dordrecht: Reidel.

Braine, M. (1963). The ontogeny of english phrase structure. *Language*, 39, 1–14.

Brandt-Kobele, O.-C., & Höhle, B. (2010). What asymmetries within comprehension reveal about asymmetries between comprehension and production: The case of verb inflection in language acquisition. *Lingua*, 120(8), 1910–1925.

Brattico, P. (2021). Finnish word order: Does comprehension matter? *Nordic Journal of Linguistics*, 44(1), 38–70. (DOI: https://doi.org/10.1017/S0332586520000098)

Brown, R. (1973). *A first language. the early stages.* Cambridge, MA: Harvard University Press.

Candan, A., Küntay, A., Yeh, Y.-C., Cheung, H., Wagner, L., & Naigles, L. (2012). Language and age effects in children's processing of word order. *Cognitive Development*, 27, 205–221.

Cauvet, E., Limissuri, R., Millotte, S., Skoruppa, K., Cabrol, D., & Christophe, A. (2014). Function words constrain on-line recognition of verbs and nouns in French 18-month-olds. *Language Learning and Development*, 10(1), 1-18.

Chabris, C., Lee, J., Benjamin, D., Beauchamp, J., Glaeser, E., Borst, G., ... Laibson, D. (2013). Why it is hard to find genes associated with social science traits: Theoretical and empirical considerations. *American Journal of Public Health*, 103(Suppl 1), S152–S166.

Chan, A., Lieven, E., & Tomasello, M. (2009). Children's understanding of the agent-patient relations in the transitive construction: Cross-linguistic comparisons between Cantonese, German, and English. *Cognitive Linguistics*, 20(2), 267–300.

Chang, C. (1998). The development of autonomy in preschool Mandarin Chinese-speaking children's play narratives. *Narrative Inquiry*, 8(1), 77–111.

Chang, F., Dell, G., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113, 234–272.

Chang, F., Kobayashi, T., & Amano, S. (2009). Social factors in the acquisition of a new word order. *First Language*, 29(4), 427–445.

Chang, H.-W. (1992). The acquisition of Chinese syntax. In H.-C. Chen & O.-J. Tzeng (Eds.), *Language processing in Chinese* (p. 175-206). Amsterdam: North Holland.

Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.

Chen, J. (2008). *The acquisition of verb compounding in Mandarin* (Doctoral dissertation). The Max-Planck Institute for Psycholinguistics & Free University of Amsterdam.

Cheung, H., Chang, C.-J., Ko, H., & Tsay, J. (2011). *Taiwan Corpus of Child Mandarin (TCCM).* (NSC 96-2420-H-002-030)

Childers, J., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the english transitive construction. *Developmental Psychology*, 37(6), 739–748.

Chomsky, N. (1959). Verbal behaviour. *Language*, 35(1), 26–58.

Chomsky, N. (1964). *Current issues in linguistic theory.* The Hague: Mouton.

Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar* (p. 184-221). Massachusetts: Ginn, Waltham.

Chomsky, N. (1980). *Rules and representations.* New York: Oxford University Press.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use.* New York: Praeger.

Chomsky, N. (1994). Bare phrase structure. In G. Webelhuth (Ed.), *Government and binding and the minimalist program* (p. 385-439). Oxford: Blackwell.

Chomsky, N. (1995). *The minimalist program.* Cambridge, MA: MIT Press.

Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti (Ed.), *Structures and beyond* (p. 104-131). Oxford: Oxford University Press.

Chomsky, N. (2010). Some simple evo-devo theses. How true might they be for language. In R. Larson, V. Deprez, & H. Yamakido (Eds.), *The evolution of human language* (p. 45-62). Cambridge: Cambridge University Press.

Chomsky, N. (2013). Problems of projection. *Lingua*, 130, 33–49.

Chomsky, N. (2015). Some core contested concepts. *Journal of Psycholinguistic Research*, 44(1), 91–104.

Chomsky, N. (2019). *The UCLA lectures.* https://linguistics.ucla.edu/noam-chomsky/. (Accessed: 25/02/2021)

Chomsky, N., Gallego, A., & Ott, D. (2019). Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics*, Special Issue, 229–261. (DOI: https://doi.org/10.5565/rev/catjl.288)

Chu, C. (1979). The word order change in Chinese and its problems. In C. Chu (Ed.), *Yuyanxue lunji [on linguistics]* (p. 87-121). Taipei: Crane Books.

Cinque, G. (1999). *Adverbs and functional heads: A cross-linguistic perspective.* New York: Oxford University Press.

Cinque, G. (2006). *Restructuring and functional head: The cartography of syntactic structure, Volume 4.* New York: Oxford University Press.

Clark, E. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (p. 1-33). New Jersey: Lawrence Erlbaum.

Coppola, M., & Newport, E. L. (2005). Grammatical subjects in home sign: Abstract linguistic structure in adult primary gesture systems without linguistic input. *Proceedings of the National Academy of Sciences*, 102(52), 19246–19253.

Crain, S., Thornton, R., & Murasugi, K. (2009). Capturing the evasive passive. *Language Acquisition*, 16(2), 123–133.

Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern-day wild child.* New York: Academic Press.

Dalrymple, K., Manner, M., Harmelink, K., Teska, E., & Elison, J. (2018). An examination of recording accuracy and precision from eye tracking data from toddlerhood to adulthood. *Frontiers in Psychology*, 9, 1–12. (DOI: https://doi.org/10.3389/fpsyg.2018.00803)

Dąbrowska, E. (2015). What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology*, 6, 852. (DOI: https://doi.org/10.3389/fpsyg.2015.00852)

Dąbrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences*, 28, 604–615.

De Cat, C. (2009). Experimental evidence for preschoolers' mastery of "topic". *Language Acquisition*, 16, 224–239.

Demuth, K. (1996). Collecting spontaneous production data. In D.McDaniel, C.McKee, & H.S.Cairns (Eds.), *Methods for assessing children's syntax* (p. 3-22). Cambridge, MA: MIT Press.

Demuth, K., & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language*, 35(1), 99–127. (DOI: https://doi.org/10.1017/S0305000907008276)

Deng, X., & Yip, V. (2018). A multimedia corpus of child Mandarin: The Tong corpus. *Journal of Chinese Linguistics*, 46(1), 69–92.

Dodson, K., & Tomasello, M. (1998). Acquiring the transitive construction in english: The role of animacy and pronouns. *Journal of Child Language*, 25, 555–574.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67(3), 547–619.

Dryer, M. (1992). The Greenbergian word order correlations. *Language*, 68, 81–138.

Dryer, M. (2013). Order of subject, object and verb. In M. Dryer & M. Haspelmath (Eds.), *The world atlas of language structure online.* Leipzig: Max Planck Institute for evolutionary anthropology. https://wals.info/feature/81A#2/18.0/152.9. (Accessed 09/05/2020.)

Dye, C., Kedar, Y., & Lust, B. (2019). From lexical to functional categories: New foundations for the study of language development. *First Language*, 39(1), 9–32.

Eberhard, D., Simons, G., & Fennig, C. (2021). *Ethnologue: Languages of the world.* http://www.ethnologue.com. (Accessed: 06/03/2021)

Erbaugh, M. (1982). *Coming to order: Natural selection and the origin of syntax in the Mandarin-speaking child* (Doctoral dissertation). University of California, Berkeley.

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59(5), 1–173.

Franck, J., & Lassotta, R. (2012). Revisiting evidence for lexicalized word order in young children. *Lingua*, 122, 92-106.

Franck, J., Millotte, S., & Lassotta, R. (2011). Early word order representations: Novel arguments against old contradictions. *Language Acquisition*, 18(2), 121–135.

Franck, J., Millotte, S., Posada, A., & Rizzi, L. (2013). Abstract knowledge of word order by 19 months: An eye-tracking study. *Applied Psycholinguistics*, 34, 323–336.

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, 119, 67–88.

Friedmann, N., & Costa, J. (2011). Acquisition of SV and VS order in Hebrew, European Portuguese, Palestinian Arabic, and Spanish. *Language Acquisition*, 18(1), 1–38.

Gagliardi, A., Mease, T., & Lidz, J. (2016). Discontinuous development in the acquisition of filler-gap dependencies: Evidence from 15- and 20-month-olds. *Language Acquisition*, 23(3), 234–260.

Gao, Q. (1997). Resultative verb compounds and *ba* construction in Chinese. *Journal of Chinese Linguistics*, 25(1), 84–130.

Gavarró, A., Leela, M., Rizzi, L., & Franck, J. (2015). Knowledge of the OV parameter setting at 19 months: Evidence from Hindi-Urdu. *Lingua*, 154, 27–34.

Gavarró, A., Lewandowski, W., & Markova, A. (2010). An approach to multiple interrogatives in child Bulgarian and Polish. In J.Costa, A.Castro, M.Lobo, & F.Pratas (Eds.), *Language acquisition and development. proceedings of generative approaches to language acquisition 2009* (p. 170-181). Newcastle: Cambridge Scholars Publishing.

Gerken, L., Jusczyk, P., & Mandel, D. (1994). When prosody fails to cue syntactic structure: Nine-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition*, 51, 237–265.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science*, 17(8), 684–691. (DOI: https://doi.org/10.1111/j.1467-9280.2006.01767.x)

Gibson, E., & Wexler, K. (1994). Triggers. *Linguistic Inquiry*, 25(3), 407–454.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1, 3–55.

Gleitman, L., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. (2005). Hard words. *Language Learning and Development*, 1, 23-64.

Gleitman, L., January, D., Nappa, R., & Trueswell, J. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, 57, 544-569.

Gleitman, L., Liberman, M., McLemore, C., & Partee, B. (2019). The impossibility of language acquisition (and how they do it). *Annual Review of Linguistics*, 5, 1–24.

Goldin-Meadow, S. (2003). *The resilience of language.* New York: Psychology Press.

Golinkoff, R., Hirsh-Pasek, K., Cauley, K., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14, 23–45.

Golinkoff, R., Ma, W., Song, L., & Hirsh-Pasek, K. (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned? *Perspectives on psychological science*, 8(3), 316-339.

Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (Ed.), *Universals of human language* (p. 73-113). Cambridge, MA: MIT Press.

Guasti, M.-T. (2002). *Language acquisition: The growth of grammar.* Cambridge, MA: MIT Press.

Guasti, M.-T., Gavarró, A., De Lange, J., & Caprin, C. (2008). Article omission across child languages. *Language Acquisition*, 15(2), 89–119.

Han, Z. (2009). *Xin hua zi dian [Xinhua dictionary]*. Beijing: The Commercial Press.

Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese. *Behavior Research Methods*, 40(3), 728–733.

Harris, F., & Flora, J. (1982). Children's use of *get* passives. *Journal of Psycholinguistic Research*, 11(4), 297–311.

Hartshorne, J., Tenenbaum, J., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, 177, 263–277. (DOI: https://doi.org/10.1016/j.cognition.2018.04.007)

He, A.-X., & Lidz, J. (2017). Verb learning in 14- and 18-month-old English-learning infants. *Language Learning and Development*, 13(3), 335–356.

Hessels, R., Andersson, R., Hooge, I., Nyström, M., & Kemner, C. (2015). Consequences of eye color, positioning, and head movement for eye-tracking data quality in infant research. *Infancy*, 20, 601-633.

Hessels, R., Cornelissen, T., Kemner, C., & Hooge, I. (2015). Qualitative tests of remote eyetracker recovery and performance during head rotation. *Behavior Research Methods*, 47, 848–859.

Hessels, R., & Hooge, I. (2019). Eye tracking in developmental cognitive neuroscience- The good, the bad and the ugly. *Developmental Cognitive Neuroscience*, 40, 100710.

Hirsh-Pasek, K., & Golinkoff, R. (1996). *The origins of grammar*. Cambridge, MA: Cambridge University Press.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. New York: Oxford University Press.

Horgan, D. (1978). The development of the full passive. *Journal of Child Language*, 5, 65–80.

Hsu, D. (2014). Mandarin-speaking three-year-olds' demonstration of productive knowledge of syntax: Evidence from syntactic productivity and structural priming with the SVO-ba alternation. *Journal of Child Language*, 41, 1115-1146.

Hsu, D. (2018). Children's syntactic representation of the transitive constructions in Mandarin Chinese. *PLoS ONE*, 13(11). (DOI: https://doi.org/10.1371/journal.pone.0206788)

Hsu, N., Rispoli, M., & Hadley, P. (2019). Mastering the grammar of complex events: Evidence from Mandarin resultative verb compounds. *Language Learning and Development*, 15(3), 255–277.

Huang, C.-T. (1982). *Logical relations in Chinese and the theory of grammar* (Doctoral dissertation). Massachusetts Institute of Technology.

Huang, C.-T. (1984a). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15(4), 531–574.

Huang, C.-T. (1984b). Phrase structure, lexical integrity, and Chinese compounds. *Journal of Chinese Language Teachers Association*, 19, 53–78.

Huang, C.-T. (1989). Pro-drop in Chinese: A generalized control theory. In O. Jaeggli & K. Safir (Eds.), *The null subject parameter* (p. 185-214). Dordrecht: Kluwer Academic Publishers.

Huang, C.-T., Li, A.-H., & Li, Y. (2009). *The syntax of Chinese*. New York: Cambridge University Press.

Huang, C.-T., & Roberts, I. (2017). Principles and parameters of Universal Grammar. In I.Roberts (Ed.), *The Oxford handbook of Universal Grammar* (p. 307-354). Oxford: Oxford University Press.

Huang, S. (1978). Historical change of prepositions and emergence of SOV order. *Journal of Chinese Linguistics*, 6, 212–242.

Huang, Y., & Yang, S. (2004). The L2 acquisition of the Chinese Ba-construction. *Chinese Teaching in the World*, 67(1), 49–59.

Hunsicker, D., & Goldin-Meadow, S. (2012). Hierarchical structure in a self-created communication system: Building nominal constituents in homesign. *Language*, 88(4), 732–763.

Hyams, N., Ntelitheos, D., & Manorohanta, C. (2006). Acquisition of the Malagasy voicing system: Implications for the adult grammar. *Natural Language & Linguistic Theory*, 24(4), 1049–1092.

Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: Evidence form German and French infants. *Infant Behavior & Development*, 32, 262–274.

Höhle, B., Weissenborn, J., Klefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy*, 5(3), 341-353.

Jardri, R., Houfflin-Debarge, V., Delion, P., Pruvo, J.-P., Thomas, P., & Pins, D. (2012). Assessing fetal response to maternal speech using a noninvasive functional brain imaging technique. *International Journal of Developmental Neuroscience*, 30(2), 159–161.

Jheng, W.-C. (2012). *The high applicative analysis of the BA construction: Evidence from GEI.* Paper presented at the 2012 National Conference on Linguistics, Tainan.

Jin, K., & Fisher, C. (2014). Early evidence for syntactic bootstrapping: 15-month-olds use sentence structure in verb learning. In W. Orman & M. Valleau (Eds.), *BUCLD 38 Online Proceedings Supplement.* Boston, MA: Cascadilla Press. https://http://www.bu.edu/bucld/supplementvol38/. (Accessed 02/06/2020.)

Kayne, R. (1994). *The antisymmetry of syntax.* Cambridge, MA: MIT Press.

Kayne, R. (2018). *The place of linear order in the language faculty.* https://ling.auf.net/lingbuzz/003820. (Accessed: 20/03/2021)

Kedar, Y., Casasola, M., Lust, B., & Parmet, Y. (2017). Little words, big impact: Determiners begin to bootstrap reference by 12 months. *Language Learning and Development*, 13(3), 317–334.

Keidel, A., Leela, M., Zhu, J., & Kunnari, S. (2021). *A study on infant word order in Finnish. Does Case make a difference?* Paper presented at the 30 Colloquium of Generative Grammar, Barcelona [Online].

Kisilevsky, B., Hains, S., Brown, C., Lee, C., Cowperthwaite, B., Stutzman, S., ... Wang, Z. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development*, 32(1), 59–71.

Kuhl, P. K., Ramírez, R. R., Bosseler, A., Lin, J. F., & Imada, T. (2014). Infants' brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences*, 111(31), 11238–11245.

Kuo, P. (2015). The components of sideward movement in the verb copying construction in Mandarin Chinese. *Studies in Chinese Linguistics*, 36(1), 36–58. (DOI: https://doi.org/10.1515/scl-2015-0003)

Lai, Y.-H. (2018). *Processing non-canonical structures in Cantonese.* Talk given at the Goethe-University Frankfurt (GK) colloquium. Goethe-University Frankfurt. Nov. 13th.

Lassotta, R. (2021). *To be or not to be adultlike in syntax: An experimental study of language acquisition and processing in children* (Doctoral dissertation). Université de Genève.

Lassotta, R., Omaki, A., & Franck, J. (2014). *Abstract knowledge of non-canonical word order by age 2.* Paper presented at the Boston University Conference on Language Development 36, Boston.

Lee, C., & Cho, S. (2009). Acquisition of the subject and topic nominals and markers in the spontaneous speech of young children in korean. In C.Lee, G.B.Simpson, & Y.Kim (Eds.), *The handbook of east asian psycholinguistics, vol.iii* (p. 23-33). Cambridge: Cambridge University Press.

Leela, M. (2016). *Early acquisition of word order: Evidence from Hindi-Urdu and Malayalam* (Doctoral dissertation). Universitat Autònoma de Barcelona.

Legate, J., & Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition*, 14(3), 315–344.

Lenneberg, E. (1967). *Biological foundations of language.* New York: Wiley.

Li, A.-H. (1985). *Abstract case in Mandarin Chinese* (Doctoral dissertation). University of Southern California, Los Angeles.

Li, A.-H. (1990). *Order and constituency in Mandarin Chinese.* Dordrecht: Kluwer Academic Publishers.

Li, A.-H. (2006). Chinese *ba*. In M. Everaert & H. van Riemsdijk (Eds.), *Blackwell companion to syntax: Volume I* (p. 374-468). Oxford: Blackwell.

Li, C.-N., & Thompson, S. (1975). The semantic function of word order: A case study in Mandarin. In C.-N. Li (Ed.), *Word order and word order change* (p. 163-186). Austin, Texas: University of Texas Press.

Li, C.-N., & Thompson, S. (1976). Subject and topic: A new typology of language. In C.-N. Li (Ed.), *Subject and topic* (p. 457-489). New York: Academic Press.

Li, C.-N., & Thompson, S. (1981). *Mandarin Chinese: A functional reference grammar.* Berkeley, CA: University of California Press.

Li, P., Bates, E., & MacWhinney, B. (1993). Processing a language without inflections: A reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language*, 32, 169–192.

Lidz, J., Gleitman, H., & Gleitman, L. (2001). *Kidz in the hood: Syntactic bootstrapping and the mental lexicon.* Philadelphia, PA: University of Pennsylvania.

Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, 20(3), 481–507.

Liu, F. (1997). An aspectual analysis of Ba. *Journal of East Asian Linguistics*, 6, 51–99.

Lu, J. (1980). Hanyu kouyu jufali de yiwei xianxiang [Discolation in the syntax of colloquial Mandarin Chinese]. *Zhongguo Yuwen*, 1, 28–41.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear & Hearing*, 19, 1–36.

Lü, S. (1980). *Xiandai hanyu babai ci[800 words of contemporary chinese].* Beijing: Shangwu Yinshuguan.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd Edition ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (2015). *The CHILDES project: Tools for analyzing talk.* http://childes.talkbank.org/manuals/chat.pdf.

Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J., ... Wallois, F. (2013). Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proceedings of the National Academy of Sciences*, 110(12), 4846–4851.

Mampe, B., Friederici, A., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology*, 9(23), 1994-1997.

Marchetto, E., & Bonatti, L. (2015). Finding words and word structure in artificial speech: The development of infants' sensitivity to morphosyntactic regularities. *Journal of Child Language*, 42(4), 873–902. (DOI: https://doi.org/10.1017/S0305000914000452)

Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we Can trust.* Random House: Vintage.

Marácz, L. (1989). *Asymmetries in Hungarian* (Doctoral dissertation). University of Groningen.

Matsuo, A., Kita, S., Shinya, Y., Wood, G., & Naigles, L. (2012). Japanese two-year-olds use morphosyntax to learn novel verb meanings. *Journal of Child Language*, 39, 637–663.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development*, 20, 121–136.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2007). French children's use and correction of weird word orders: a constructivist account. *Journal of Child Language*, 34(2), 381–409.

Mayberry, R., & Eichen, E. (1991). The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language*, 30, 486–512.

McNeill, D. (1966). Developmental psycholinguistics. In F. Smith & G. Miller (Eds.), *The Genesis of Language: A Psycholinguistic Approach.* Cambridge, MA: MIT Press.

Mendola, J., Conner, I., Roy, A., & Chan, S. (2005). Voxel-based analysis of MRI detects abnormal visual cortex in children and adults with amblyopia. *Human Brain Mapping*, 25(2), 222–236.

Messenger, K., Branigan, H. P., & McLean, J. F. (2012). Is children's acquisition of the passive a staged process? Evidence from six- and nine-year-olds' production of passives. *Journal of Child Language*, 39(5), 991–1016.

Messenger, K., Yuan, S., & Fisher, C. (2015). Learning verb syntax via listening: New evidence from 22-month-olds. *Language Learning and Development*, 11(4), 356–368.

Miao, X. (1981). Word order and semantic strategies in Chinese sentence comprehension. *International Journal of Psycholinguistics*, 8, 109–122.

Miyata, S. (2004). *Aki corpus.* https://childes.talkbank.org/access/Japanese/Miyata.html. (Pittsburgh, PA: TalkBank)

Musso, M., Moro, A., Glauche, V., Rijnyjes, M., Reichenbach, J., Büchel, C., & Weiller, C. (2003). Broca's area and the language instinct. *Nature Neuroscience*, 6(7), 774–781.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17, 357–374. (DOI: https://doi.org/10.1017/S0305000900013817)

Naigles, L. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, 58, 221–251. (DOI: https://doi.org/10.1016/0010-0277(95)00681-8)

Naigles, L. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition*, 86(2), 157–199. (DOI: https://doi.org/10.1016/s0010-0277(02)00177-4)

Naigles, L., & Kako, E. (1993). First contact in verb acquisition: defining a role for syntax. *Child Development*, 64, 1665–1687. (DOI: https://doi.org/10.1111/j.1467-8624.1993.tb04206.x)

Nappa, R., Wessell, A., McEldoon, K. L., Gleitman, L., & Trueswell, J. (2009). Use of speakers's gaze and syntax in verb learning. *Language Learning and Development*, 5, 203–234.

Nespor, M., Guasti, M.-T., & Christophe, A. (1996). Selecting word order: The rhythmic activation principle. In U. Kleinhenz (Ed.), *Interfaces in phonology* (p. 1-26). Berlin: Akademie Verlag.

Newport, E. L. (2020). Children and adults as language learners: Rules, variation, and maturational change. *Topics in Cognitive Science*, 12(1), 153–169.

Omaki, A., Lassotta, R., Kobayashi, T., Rizzi, L., & Franck, J. (2012). Delay of word order in Japanese? Evidence from a preferential looking study with 19 and 30-month-old children. In N. Miyake, D. Peebles, &

R. Cooper (Eds.), *Proceedings of 34th Annual Meeting of the Cognitive Science Society* (p. 2826). Austin, TX: Cognitive Science Society.

Oviatt, S. (1980). The emerging ability to comprehend language: An experimental approach. *Child Development*, 51, 97–106.

Packard, J. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge: Cambridge University Press.

Patterson, F. (1978). The gestures of a gorilla: Language acquisition in another pongid. *Brain and Language*, 5, 56–71.

Patterson, F., & Cohn, R. (1990). Language acquisition by a lowland gorilla: Koko's first ten years of vocabulary development. *Word*, 41(2), 97–143.

Perkins, L., & Lidz, J. (2020). Filler-gap dependency comprehension at 15 months: The role of vocabulary. *Language Acquisition*, 27(1), 98–115.

Petitto, L. (2005). How the brain begets language. In J. McGolvray (Ed.), *The Chomsky Reader* (p. 85-101). Cambridge: Cambridge University Press.

Piatelli-Palmarini, M. (1980). *Language and Learning: The Debate Between Jean Piaget and Noam Chomsky*. Cambridge, MA: Harvard University Press.

Pierce, A. (1989). *On the emergence of syntax: A crosslinguistic study* (Doctoral dissertation). Massachusetts Institute of Technology.

Pierce, A. (1992). *Language acquisition and syntactic theory: A comparative analysis of French and English child grammar*. Dordrecht: Kluwer.

Pine, J., & Lieven, E. (1993). Reanalysing rote-learned phrases: Individual differences in the transition to multi-word speech. *Journal of Child Language*, 20(3), 551–571.

Pine, J., & Lieven, E. (1997). Slot and frame patterns in the development of the determiner category. *Applied Psycholinguistics*, 18(2), 123–138.

Poeppel, D., & Wexler, K. (1993). The full competence hypothesis of clause structure in early German. *Language*, 69, 365-424.

Pons, F., Lewkowicz, D. J., Soto-Faraco, S., & Sebastian-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences*, 106(26), 10598–10602.

R Development Core Team. (2015). *The r project for statistical computing.* http://www.r-project.org. (Accessed: 05/06/2020)

R Development Core Team. (2017). *R: A language environment for statistical computing.* http://www.r-project.org. (Accessed: 05/06/2020)

Ramchand, G. (2008). *Verb meaning and the lexicon.* Cambridge: Cambridge University Press.

Rice, M. (2012). Toward epigenetic and gene regulation models of specific language impairment: Looking for links among growth, genes, and impairments. *Journal of Neurodevelopmental Disorders*, 4(27), 1–27.

Rice, M., Haney, K., & Wexler, K. (1998). Family histories of children with SLI who show extended optional infinitives. *Journal of Speech and Hearing Research*, 41(2), 419–432.

Richards, M. (2008). Two kinds of variation in a minimalist system. In F. Heck, G. Müller, & J. Trommer (Eds.), *Varieties of competition* (p. 133-162). Universität Leipzig: Linguistische Arbeits Berichte 87.

Rizzi, L. (2001). On the position "Int(errogative)" in the Left Periphery of the Clause. In G. Cinque & G. Salvi (Eds.), *Current studies in Italian syntax: Essays offered to Lorenzo Renzi* (p. 287-296). New York: Elsevier.

Roberts, I. (2016). *Null subject parameters and determiner parameters.* Paper presented at the 1st SynCart workshop: From maps to principles: Syntactic cartography and locality, Chiusi.

Roberts, I., & Holmberg, A. (2010). Introduction: Parameters in minimalist theory. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan (Eds.), *Parametric variation: Null subjects in minimalist theory* (p. 1-56). Cambridge: Cambridge University Press.

Robertson, E., Shi, R., & Melançon, A. (2012). Toddlers use the number feature in determiners during online noun comprehension. *Child Development*, 83, 2007–2018.

Roeper, T. (2007). What frequency can do and what it can't. In I.Gülzow & N.Gagarina (Eds.), *Frequency effects in language acquisition: Defining the limits of frequency as an explanatory concept* (p. 23-50). Berlin: Mouton de Gruyter.

Rowland, C. F., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition*, 125(1), 49–63.

Russell, B. (1948). *Human knowledge: Its scope and limits.* New York: Simon & Schuster.

Santelmann, L., & Jusczyk, P. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, 69, 105–134.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: Lexical and structural priming of syntactic constructions in young children. *Developmental Science*, 6(5), 557–567.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4-year-olds. *Language Learning and Development*, 2, 27–50.

Seidl, A., Hollich, G., & Jusczyk, P. W. (2003). Early understanding of subject and object wh-questions. *Infancy*, 4(3), 423–436.

Senghas, A. (1995). *Children's contribution to the birth of Nicaraguan Sign Language* (Doctoral dissertation). Massachusetts Institute of Technology.

Senghas, A., & Coppola, M. (2001). Children Creating Language: How Nicaraguan Sign Language Acquired a Spatial Grammar. *Psychological Science*, 12(4), 323–328.

Senghas, A., Coppola, M., Newport, E., & Supalla, T. (1997). Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the 21st Boston University Conference on Language Development.* Boston, MA: Cascadilla Press.

Shimpi, P., Gámez, P., Huttenlocher, J., & Vasilyeva, M. (2007). Syntactic priming in 3- and 4-year-old children: Evidence for abstract representations of transitive and dative forms. *Developmental Psychology*, 43(6), 1334–1346.

Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: Tthe acquisition of American Sign Language from inconsistent input. *Cognitive Psychology*, 49, 370–407.

Smith, N., & Tsimpli, I. (1995). *The mind of a savant.* Oxford: Blackwell.

Snyder, W., & Hyams, N. (2015). Minimality effects in children's passives. In E. di Domenico, C. Hamann, & S. Matteini (Eds.), *Structures,* *strategies and beyond: Studies in honour of Adriana Belletti* (p. 343-368). Amsterdam: John Benjamins.

Stowell, T. (1981). *Origins of phrase structure* (Doctoral dissertation). Massachusetts Institute of Technology.

Sugisaki, K. (2008). Early acquisition of basic word order in Japanese. *Language Acquisition*, 15(3), 183–191. (DOI: https://doi.org/10.1080/10489220802142441)

Sun, C., & Givon, T. (1985). On the so-called sov word order in Mandarin Chinese: A quantified text study and its implications. *Language*, 61, 329–351.

Sun, Y. (2019). A cartographic analysis of the syntactic structure of Mandarin Ba. *Studies in Chinese Linguistics*, 39(2), 128–154. (DOI: https://doi.org/10.2478/scl-2018-0005)

Sybesma, R. (1992). *Causatives and accomplishments: The case of Chinese ba.* (Doctoral dissertation). Leiden University.

Tai, J.-Y. (1973). Chinese as a SOV language. In *Proceedings of the 9th Regional Meeting. Chicago Linguistic Society* (pp. 659–671). Illinois.

Tardif, T. (2015). Early vocabulary in Chinese-speaking children. In W.-Y. Wang & C. Sun (Eds.), *The Oxford handbook of Chinese linguistics* (p. 641-653). New York: Oxford University Press.

Terrace, H., Petitto, L., Sanders, R., & Bever, T. (1979). Can an ape create a sentence? *Science*, 206(4421), 891–902.

Thornton, R. (2008). Why continuity. *Natural Language and Linguistic Theory*, 26, 107-146.

Thornton, R. (2016). Acquisition of questions. In J. Lidz, W. Snyder, & J. Pater (Eds.), *The Oxford handbook of Developmental Linguistics* (p. 310-340). Oxford: Oxford University Press.

Thornton, R., & Tesan, G. (2007). Categorical acquisition: Parameter-setting in Universal Grammar. *Biolinguistics*, 1, 49–98.

Thorpe, W. (1958). The learning of song patterns by birds, with especial reference to the song of the Chaffinch Fringilla coelebs. *Ibis*, 100, 535–570. (DOI: https://doi.org/10.1111/j.1474-919X.1958.tb07960.x)

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74(3), 209–253. (DOI: https://doi.org/10.1016/S0010-0277(99)00069-4)

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomioka, S. (2003). The semantics of Japanese null pronouns and its cross-linguistic implications. In K. Schwabe & S. Winkler (Eds.), *The interfaces: Deriving and interpreting omitted structures* (p. 321-340). Amsterdam: John Benjamins.

Travis, L. (1984). *Parameters and effects of word order variation* (Doctoral dissertation). Massachusetts Institute of Technology.

Tsai, D.-T. (2008). Object specificity in Chinese: A view from the vP periphery. *The Linguistic Review*, 25, 479–502.

Tsao, F. (1987). A topic-comment approach to the ba construction. *Journal of Chinese Linguistics*, 15, 1-54.

Tse, K. (2020). Differential object marking: Nominal and verbal parameters. *Proceedings of the Linguistic Society of America*, 5(1). (DOI: https://doi.org/10.3765/plsa.v5i1.4748)

Van Bergen, G. (2006). *To ba or not to ba: Differential object marking in Chinese* (M.A. Thesis). Radboud University Nijmegen.

Van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science*, 12, 419–425.

Vilkuna, M. (1995). Discourse configurationality in Finnish. In K. Kiss (Ed.), *Discourse Configurational Languages* (p. 244-268). New York: Oxford University Press.

Wang, H. (1957). *Ba zi ju he bei zi ju [On the ba sentences and bei sentences]*. Shanghai: Shanghai Jiaoyu Chubanshe.

Wang, L. (1945). *Zhongguo yufa lilun [Theories of Chinese syntax]*. Beijing: Zhonghua Press.

Wang, L. (1947). *Modern chinese grammar*. Shanghai: Zhonghua Press.

Wang, M. (1987). *Transitivity and the Ba-construction in Mandarin* (Doctoral dissertation). Boston University.

Wang, Q., Lillo-Martin, D., Best, C., & Levitt, A. (1992). Null subject versus null object: Some evidence from the acquisition of Chinese and English. *Language Acquisition*, 2(3), 221–254.

Wells, C. (1981). *Learning through interaction: The study of language development*. Cambridge: Cambridge University Press.

Wen, X. (2012). A daunting task? The acquisition of the Chinese *ba*-construction by nonnative speakers of Chinese. *Journal of Chinese Linguistics*, 40(1), 216–240.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation for the optional infinitive. *Lingua*, 106, 23–79.

Wexler, K. (2004). Theory of phasal development: Perfection in child grammar. *MIT Working Papers in Linguistics*, 48, 159–209.

Wexler, K. (2011). Grammatical computation in the optional infinitive stage. In J. de Villiers & T. Roeper (Eds.), *Handbook of generative approaches to language acquisition* (p. 53-118). Dordrecht: Springer.

Yang, C. (2002). *Knowledge and learning in natural language.* Oxford: Oxford University Press.

Yang, C. (2004). Universal Grammar, statistics, or both? *Trends in Cognitive Science*, 8(10), 541–456.

Yang, C. (2005). On productivity. *Linguistic Variation Yearbook*, 5(1), 265–302.

Yang, C. (2013). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences*, 110(16), 6324–6327.

Yang, C. (2016). *The price of productivity.* Cambridge, MA: MIT Press.

Yang, X., Shi, R., & Xu, K. (2018). Grammatical aspect in early child Mandarin: Evidence from a preferential looking experiment. *Journal of Psycholinguistic Research*, 47(4), 1–20. (DOI: https://doi.org/10.1007/s10936-018-9590-7)

Yeh, M.-Y. (2015). *The role of maternal input in early word order acquisition: The case of Mandarin Chinese* (Doctoral dissertation). University of Connecticut.

Yokoyama, M., & Miyata, S. (2017). *Yokoyama corpus.* https://childes.talkbank.org/access/Japanese/Yokoyama.html. (Pittsburgh, PA: TalkBank)

Yu, S., & Tamaoka, K. (2018). Age-related differences in the acceptability of non-canonical word orders in Mandarin Chinese. *Lingua Sinica*, 4(3). (DOI: https://doi.org/10.1186/s40655-018-0035-x)

Yu, X., & Zhang, H. (2019). Word order as interface between syntax and pragmatics: The case of identifying topics in mixed case-marking patterns in Mandarin Chinese. *Lingua*, 223, 2–28.

Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, 20, 619–626.

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: simple structural cues to verb meaning. *Child Development*, 83, 1382–1399.

Zhan, L. (2018). Using eye movements recorded in the visual word paradigm to explore the online processing of spoken language. *Journal of Visualized Experiments*, 140. (DOI: https://doi.org/10.3791/58086)

Zhang, Z., Shi, R., & Li, A. (2015). Grammatical categorization in Mandarin-Chinese-learning infants. *Language Acquisition*, 22(1), 104-115.

Zhou, J. (2001). *Pragmatic development of Mandarin speaking young children: From 14 months to 32 months* (Doctoral dissertation). The University of Hong Kong.

Zou, K. (1995). *The syntax of the Chinese ba-construction and verb compounds: A morpho-syntactic analysis* (Doctoral dissertation). University of California, Irvine.

A Appendix

A.1 Five infants' individual production of null/overt subjects at 1;8

(1) a. MOT: Bi yao bu yao?
pen want NEG want
'Do you want a pen?'

b. CHI: Yao.
want
'(I) want (it).'

(Qinlong, 1;8)

(2) a. MOT: Wo-men lai tuo yi-fu?
we come undress clothes
'Should we take our clothes off?'

b. CHI: Bu yao tuo yi-fu.
NEG want undress clothes
'(I) do not want to take the clothes off.'

(Haohao, 1;8)

(3) a. MOT: Ca shen-me?
wipe what
'What do you wipe?'

b. CHI: Ca fei-zao.
wipe soap
'(I) wash with soap.'

(Wenwen, 1;8)

(4) a. Zhe ge bu yao bai hao.
this CLASS NEG want put good
'(You) do not put this back.'

b. Wo yao na na ge.
I want take that CLASS

'I want to take that.' (Jiang Weiying, 1;8)

(5) a. MOT: Xiao-ji chi shen-me?
chicken eat what
'What do chickens eat?'

b. CHI: Chi mi.
eat rice
'(They) eat rice.'

(Xue'er, 1;8)

A.2 Five infants' individual production of null/overt objects at 1;8

(1) a. MOT: Hai yao bu yao a?
still want NEG want Q
'Do you still want it?'

b. CHI: Yao.
want
'(I) want (it).'

(Qinlong, 1;8)

(2) a. MOT: Wo-men jiao shou-biao.
we call watch
'We call it watch.'

b. CHI: Bao-bao ye you.
baby also have
'I also have (one).'

(Haohao, 1;8)

(3) a. MOT: Gan shen-me?
do what
' What do you want me to do?'

b. CHI: Ba.
pull out
'(You) pull (it) out.'

(Wenwen, 1;8)

(4) a. Ma-ma qu na.
mom go get
'Mom go and get (it).'

b. Ma-ma zai hua ge qi-qiu.
mom again draw CLASS balloon
'Mom, draw a balloon again.' (Jiang Weiying, 1;8)

(5) a. MOT: Li-mian quan shi shu.
inside all is book
'There are many books inside.'

b. CHI: Na bu dao.
take NEG arrive
'(I) cannot take (them).'

(Xue'er, 1;8)

A.3 Chinese Experiment I Test Sentences

A.3.1 Grammatical condition

(1) a. Xiao-gou chei le xiao-lv.
little-dog PSEUDO-V PERF little-donkey
'The puppy V-ed the little donkey.'

b. Shi-zi chei le xiao-ma.
lion PSEUDO-V PERF little-horse
'The lion V-ed the foal.'

c. Xiao-niu chei le xiao-yang.
little-cow PSEUDO-V PERF little-sheep
'The calf V-ed the lamb.'

A.3.2 Ungrammatical condition

(2) a. * Xiao-niu shi-zi nui le.
little-cow lion PSEUDO-V PERF
'The calf the lion V-ed.'

b. * Xiao-lv xiao-gou nui le.
little-donkey little-dog PSEUDO-V PERF
'The little donkey the puppy V-ed.'

c. * Xiao-yang xiao-ma nui le.
little-sheep little-horse PSEUDO-V PERF
'The lamb the foal V-ed.'

A.4 Chinese Experiment II Test Sentences

A.4.1 SVO condition

(1) a.

Xiao-tu-zi	tuan	le	xiao-gou.
little-rabbit	PSEUDO-V	PERF	little-dog

‘The bunny V-ed the puppy.’

b.

Xiao-ya-zi	tuan	le	xiao-ma.
little-duck	PSEUDO-V	PERF	little-horse

‘The duckling V-ed the foal.’

c.

Xiao-niu	tuan	le	xiao-ma.
little-cow	PSEUDO-V	PERF	little-horse

‘The calf V-ed the foal.’

A.4.2 S*ba*OV condition

(2) a.

Xiao-tu-zi	ba	xiao-ya-zi	tuan	le.
little-rabbit	BA	little-duck	PSEUDO-V	PERF

‘The bunny V-ed the duckling.’

b.

Xiao-niu	ba	xiao-ma	tuan	le.
little-cow	BA	little-horse	PSEUDO-V	PERF

‘The calf V-ed the foal.’

c.

Xiao-yang	ba	xiao-gou	tuan	le.
little-sheep	BA	little-dog	PSEUDO-V	PERF

‘The lamb V-ed the puppy.’

A.4.3 OS*ba*OV condition

(3) a.

Xiao-ya-zi,	Xiao-tu-zi	ba	ta	tuan	le.
the-little-duck$_i$	little-rabbit	BA	it$_i$	PSEUDO-V	PERF

'The duckling$_i$, the bunny V-ed it$_i$.'

b.

Xiao-gou,	Xiao-niu	ba	ta	tuan	le.
the-little-dog$_i$	little-cow	BA	it$_i$	PSEUDO-V	PERF

'The puppy$_i$, the calf V-ed it$_i$.'

c.

Xiao-yang,	Xiao-ma	ba	ta	tuan	le.
the-little-sheep$_i$	little-horse	BA	it$_i$	PSEUDO-V	PERF

'The lamb$_i$, the foal V-ed it$_i$.'