

ARTICLE

Modeling monolingual and bilingual children's language attitudes towards variation in metropolitan France

Anna Ghimenton^{1,2}, Christophe Coupé³, Nelly Bonhomme², Jinke Song² and Vincent Arnaud^{2,4} 

¹Laboratoire Lidilem (UR 609) – UFR LLASIC, Université Grenoble Alpes, France, ²Laboratoire Dynamique du Langage, Université de Lyon, CNRS, France, ³Department of Linguistics, University of Hong Kong, Hong Kong and ⁴Département des arts, des lettres et du langage, Université du Québec à Chicoutimi, QC, Canada
Corresponding author: Vincent Arnaud; Email: vincent.arnaud@uqac.ca

(Received 19 February 2024; revised 11 March 2025; accepted 12 March 2025)

Abstract

This study investigates four factors (age, sex, SES, and bilingualism) influencing children's language attitude (LA) development. We examine LAs in monolingual (N = 46) and bilingual (N = 71) children (59–143 months) living in France using a matched guise experiment where the children evaluated normative and non-normative variants of five linguistic constructions in French. Using a mixed-effects model, we show that children's preferences for normative variants increase with age, and each linguistic construction documented is subject to different attitudinal timeframes. The probabilities of preferring the normative variants are significantly higher for monolingual girls than for bilingual girls. Whilst lower-class and upper-class children's LAs are similar, low-to-middle-class children's responses are more random, which may illustrate the potential effects of linguistic insecurity. We discuss how the children's construction of the sociocognitive representations of linguistic variation could be explained by considering children's language exposure and experiences of socialisation.

Keywords: language attitudes; acquisition of sociolinguistic norms; bilingual/monolingual

Résumé

Cette étude examine l'influence de quatre facteurs (âge, sexe, statut social, bilinguisme) sur le développement des attitudes linguistiques (AL) des enfants. Nous examinons les AL chez des enfants français monolingues (N=46) et bilingues (N=71) (59–143 mois) en utilisant la technique du locuteur masqué au cours duquel les enfants devaient évaluer des variantes normatives et non normatives de cinq constructions linguistiques. En utilisant un modèle de régression à effets mixtes, il apparaît que les préférences des enfants pour les variantes normatives augmentent avec l'âge et que les attitudes liées à chaque construction linguistique analysée sont sujettes à des chronologies spécifiques. Les probabilités de préférer les variantes

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

normatives sont significativement plus élevées pour les filles monolingues que pour les filles bilingues. Alors que les AL des enfants des classes sociales populaires et favorisées sont similaires, les réponses des enfants des classes moyennes sont plus aléatoires, illustrant les effets potentiels d'une insécurité linguistique. Nous examinons la façon dont la construction des représentations sociocognitives de la variation linguistique chez les enfants peut être expliquée en tenant compte de leur exposition à la langue et de leurs expériences de socialisation.

1. Introduction

Attitudes express psychological evaluations that show a degree of favour or disfavour towards a specific entity (Eagly & Chaiken, 1998). Language attitudes (henceforth, LA) express positive or negative evaluations or preferences towards a linguistic stimulus and are complex, socially embedded cognitive constructs (Coronel-Molina, 2014). Language variation is an interesting venue to tap into speakers' LA. Indeed, a part of LA research has investigated how language variability is perceived with different variants representing “alternate ways of saying the same thing” (Labov, 1972:188). The use of one variant as opposed to another may depend on different factors. For example, bilingual children may choose one variant over another to accommodate to their interlocutor's language choices via pragmatic adjustments (Ghimenton *et al.*, 2013), or the variants produced may be forms of alignment in conversations (Garrod & Pickering, 2004). Children start producing variation early (see review in Sim & Post, 2024), and this reflects the process of language socialisation through their interactions with different interlocutors in their social circles, starting with their caregivers and later on with their peers.

LAs are influenced by processes of language standardization (Bouchard Ryan & Giles, 1982; Garrett, 2010) which push towards uniformity and perceived correctness (Gadet, 2007; Milroy & Gordon, 2003). Normative variants are supported by “elite” ideologies, impacting LA preferences (Bouchard Ryan & Giles, 1982). Philosopher Michel Foucault wrote extensively on the pressures towards conformity and on the power exerted by institutions on the individuals (Foucault, 1975). Schools fall under this category and are places of socialisation where the normative variants are promoted in discourses of correctness produced by teachers, yielding metalinguistic reflections on the appropriate uses of language. These discourses are undoubtedly influenced by prescriptive grammar and writing (CherVEL, 1977, 2008; Gadet, 2007). Yet school is also a context where children are socialised with their peers and are in contact with heterogeneous language practices. Nardy *et al.* (2014) showed this influence in a longitudinal study conducted in a French classroom setting investigating production and the evaluation of three linguistic variables: optional liaisons, optional deletion of post-consonantal /ʁ/ in final word position and the optional deletion of /l/ in the clitic pronoun context. The 4–5-year-olds in the sample did not converge towards the teacher's production, which contained more standard variables, but rather to the one of the most popular peers, which in that context contained more non-standard variants. Despite this, the children's evaluations remained stable throughout the length of the study. Moreover, some linguistic variants may carry systematic socio-indexical patterns in both production and perception, meaning that an individual may associate a non-linguistic attribute (gender, popularity, and intelligence) to the use of that particular linguistic variant (Ochs, 1990). For example, in French, the realisation of both adverbial elements in the construction of negation (*ne ... pas*) may be perceived as

a speech feature of an individual who has a well-paid job and a high status in society (Gadet, 2007).

Numerous studies on LA have investigated the developmental patterns underpinning the emergence of preferences towards the normative variants: In some studies from different linguistic communities, 8 years appears to be a pivotal age where children start to prefer standard variants over non-standard ones (Kinzler & DeJesus, 2013). Findings from bidialectal communities have shown a different scenario, indicating that children expressed preferences for standard variants between ages 3 and 6 (Day, 1980; Fehér, 2020; Rosenthal, 1974). Further, in a Francophone setting, Barbu et al. (2013) investigated the evolution in 2- to 6-year-old monolingual children's production and evaluation of standard and non-standard variants in metropolitan France. They found that only the 5–6-year-old children from higher socioeconomic status (SES) families preferred the standard variants. Investigating how 5- to -12-year-old children develop their socio-semiotic landscapes, Vaughn and Becker (2024) showed that the children's perception and categorisation of variation become increasingly more fine-grained with age as their language experience and socialisation widen.

Differences between studies in the sociolinguistic contexts investigated, in the nature of the linguistic variables (syntactic, lexical, and phonological), in the definition of concepts like monolingualism and bilingualism, and finally in the method designs may explain the variability in their results. In addition, very few studies, if any, have been conducted in metropolitan France where monolingual and bilingual children's LA towards linguistic variation are compared, i.e., comparing monolingual and bilingual school-going children's LA development to see if their respective language experiences, within their monolingual or bilingual socialisation contexts, influence the way their language preferences unfold over time.

This paper aims to fill this gap and reports the results from an experiment conducted in metropolitan France investigating the emergence of monolingual and bilingual children's language-based preferences towards sociolinguistic variables attested in the developmental literature available. The aim is twofolded. First, we document the factors that influence the emergence of LA towards normative and non-normative variants within the same language (French) through one explanatory regression model. Second, we examine whether LA towards language variation unfolds in similar or different ways in bilingual and monolingual children. In the next sections, we present a brief review of the literature on the socio-psychological foundations of the development of LA, reporting the results from studies on LA towards different linguistic variables and non-linguistic factors that may influence LA. We then present the research method, report the results, and discuss the factors explaining the children's LA's developmental trajectories.

1.1. A developmental approach to the emergence of attitudes towards linguistic variables

Children are sensitive to the variation in the input and show preferences towards their L1 or familiar accents until the age of 5 (Kinzler & DeJesus, 2013). Their exposure to the linguistic variants used in their language environment influences their LAs measured in terms of their preferences towards one rather than the other variant. LA towards phonological variables are by and large the most documented, in particular in psychology and sociolinguistic studies on language variation. In France, the focus has been placed primarily on judgements on large-scale diatopic or diastratic variation of French

(perceptions of two varieties of French spoken in Jura versus in Toulouse (Girard *et al.*, 2008); alternation between standard and non-standard realisations of the post-consonantal /ʁ/ (Chevrot *et al.*, 2000) or of optional liaisons (Nardy, 2008). Barbu *et al.* (2013) investigated both the production and perception of obligatory and optional liaisons in Francophone monolingual children aged between 2 and 6 years. Via a matched-guise protocol, the children had to indicate their preferences between either the normative or the non-normative variant. The results showed that as early as 3, children's judgements and productions are intimately linked, regardless of their respective social backgrounds. However, by age 5, the children's social backgrounds impacted both their productions and judgements. The authors suggest that although children are capable of evaluating different linguistic forms at an early age, they are not necessarily capable of decoding the indexical value carried by standard and non-standard liaison variants.

Kinzler and DeJesus's (2013) speaker evaluation design study comparing American monolingual Anglophone children's perceptions (aged 5–6 and 9–10) of northern (prestigious) and southern accents (non-prestigious) showed that the youngest in their sample preferred their home variety, regardless of the variants' prestige. However, the older children (9-year-olds), regardless of their geographical provenance, preferred the English variants of prestige. In particular, speakers of the Northern variety were perceived as being more intelligent than speakers with an accent from the south, who were viewed merely as "nice." By age 9, the children aligned their preferences with those found in adults (Preston, 1998), favouring the more prestigious accents rather than less prestigious ones. Investigating LA in a Francophone Belgian school context, Lafontaine (1986) tested liaisons (phonological) as well as lexical and morphological variables and showed that by age 12, preteens prefer the prestigious variants. In addition, the children's LA matched the "institutional stance" towards the local varieties: the teachers in the school manifested the same preference patterns as the 12-year-olds, notwithstanding their own effective use of local regional features, the latter being viewed negatively. In another sociolinguistic context where one prestigious variety is in contact with a regional and less prestigious one, Cremona and Bates (1977) investigated children's (6–10-year-old) preferences towards standard Italian, the prestige variety, compared to the local dialect. While the 6-year-olds in this study had equal preferences towards Italian and the local dialect, the 8-year-olds favoured the standard (100% preference rates), even if these preferences did not reflect their effective practices. Similarly, in the British context, Giles *et al.* (1983) noticed that by the time they were 10 years old, children from Bristol preferred Received Pronunciation, the prestige variety, over Welsh English.

Other findings in bidialectal contexts showed that preferences for the normative variants may emerge even earlier: Day (1980) and Rosenthal (1974) showed that children as early as 3 were already able to express language-based social awareness by preferring the prestigious variants. Kinzler *et al.* (2012) suggested that early emergence of social awareness may be due to a bidialectal or bilingual sociolinguistic context, which may awaken children's LA earlier than in monolingual contexts. Indeed, Rosenthal (1974) found that children's social awareness becomes more sophisticated between 3 and 5 years as they express preferences towards the normative, mainstream American English in contrast with African American Vernacular English. In bidialectal Austria, however, Kaiser and Kasberger's (2014) investigation showed another result based on a discrimination protocol of phonological and lexical variables: the 7–8-year-old Austrian German children in their cohort preferred the standard varieties. In another bidialectal sociolinguistic situation in Switzerland, Häcki Buhofer and Studer (1993) tapped into primary school children's preferences towards lexical variants from two varieties in the German-

speaking Canton. In their study, they found that 6-year-old children did not express any specific preferences towards any particular variety. Yet, by age 7, children preferred standard German, and from 8 years onwards, they manifested attitudes favouring the regional variety over the standard one.

The contrasting sociolinguistic contexts and the different nature of the linguistic variables tested (phonological, lexical, morphophonological, etc.) may explain the different results reported hitherto. However, considering the results altogether, it would appear that a preference towards the prestigious variables emerges between the ages of 3 and 12 years.

1.2. Socioeconomic status, bilingualism and language prestige as factors influencing children's attitudes

SES has been well documented in developmental literature on language perceptions (Day, 1980; Lambert et al., 1966; Nardy et al., 2014; Rosenthal, 1974). The way this factor has been operationalised in the literature reveals a particularly complex, socioculturally specific phenomenon. Related to the family's SES is the parents' level of instruction because of the "symbolic capital" (Bourdieu, 1979) it carries in relation to social mobility afforded via higher levels of education. The parents' professions have also been considered as pertinent criteria for the establishment of social classes (see among others Barbu et al., 2013; Barbu et al., 2014; Chevrot et al., 2000). Sometimes used as a proxy for SES, the mothers' level of education seems to have a noticeable influence on both production (Hoff, 2006; Hoff et al., 2002) and judgements and attitudes (Barbu et al., 2013; Kaiser & Kasberger, 2018). In Kaiser and Kasberger (2018), three SES groups were established depending on the mothers' degree: low (no high school diploma), middle (high school diploma) and high (university/college degree). Although this factor has been dealt with in varying ways, most of the literature available limits this factor to two categories: high or low SES, often based on the mother's profession. Yet societal trends show that mothers are increasingly involved in the workforce. Nardy (2008) indicated that in France, in 2008, 82.3% of women aged 25–49 were employed. This indicates that childcare is increasingly shared by both parents, meaning that the father's role should better be accounted for.

Being monolingual or bilingual has also been investigated as a potential factor affecting LA, yet its influence is less clear. Souza et al. (2013) compared 4–6-year-old monolingual and bilingual preferences towards non-native and accented native speech samples. They found that preferences for in-group members using a familiar language variety were shared by both monolinguals and bilinguals and that bilinguals did not manifest greater social flexibility compared to monolinguals. Byers-Heinlein et al. (2017) investigated whether 4–6-year-old monolingual and bilingual children's preferences were impacted by their bilingual status and by the mono/bilingual status of their interlocutors. Children in their sample had to choose the speaker they wanted to be friends with after listening to two verbal stimuli produced either by monolingual and bilingual adults or by two monolinguals. Monolinguals showed preferences towards monolingual speakers. Bilinguals displayed similar preferences towards monolinguals who used their dominant language.

The influence of language prestige was also discussed as a potential factor influencing LA. Yet, DeJesus et al. (2017) found no influence on preferences based on language status and found that children, monolinguals and bilinguals, preferred productions from native-accented speakers. In her investigation on status-based preferences in Hungary, Fehér

(2020) found that by the age of 6, both monolingual and bidialectal children preferred the standard variety over the local dialect.

Lambert *et al.* (1966) conducted a large-scale study (373 French Canadian girls between the ages of 9 and 18) in Canada aiming to understand at what age negative LAs emerged. In this particular context and at the time of the study, English was considered to be the most prestigious language. Children aligned their preferences with this mainstream view by the age of 12. Particularly stable were the LAs of the “upper-class” bilingual girls in private schools. Yet, this factor should be handled with caution because Anisfeld and Lambert’s (1964) study yielded different results: bilinguals did not show any preferences towards the prestigious variety, whereas monolinguals did. This factor thus needs further supporting evidence with a greater control of other (potentially confounding) factors of variation presumed to influence attitudes (SES, influence of institutions like schools, definition of bilingualism, etc.).

1.3. Research questions and hypotheses

A few gaps appear in the literature concerning LAs’ developmental agenda. Firstly, monolinguals’ LA have been studied extensively: they are generally asked to express their preferences between two variants of a linguistic variable (involving mainly phonetic variation within the same language system). Although some studies have explored more than one variable (e.g., Vaughn & Becker, 2024, for an experiment with children, and Buson *et al.* (2018) with adults), LAs from monolinguals are by far and large the most documented. Secondly, should bilingual or bidialectal individuals be involved, they are asked to categorise speakers of prestigious or less prestigious varieties in their environment (Dekker *et al.*, 2021; Kaiser & Kasberger, 2014; Rosseel *et al.*, 2018). Furthermore, comparisons between studies are difficult because these have been conducted in sociolinguistic contexts with contrasting sociopolitical ideologies and historical specificities.

In our study, we focus on different points of variation (morphosyntactic, morpho-phonological, and phonetic) within one language variety, metropolitan French, by asking primary school monolinguals and bilinguals living in France to express their preferences towards the 5 sociolinguistic constructions selected in the study (see *Methods* section for the linguistic constructions). In light of the literature presented, the following research questions were addressed and hypotheses tested.

RQ1: Do children’s LAs unfold at the same time regardless of the linguistic construction involved?

Hypothesis 1: LA literature suggests that eight is a pivotal age for the emergence of preferences for the normative variants, but LAs have mainly been documented by using phonetic variables. Considering that different linguistic levels show different developmental trajectories, children’s LAs vary depending on the linguistic construction evaluated.

RQ2: What factors explain the variation in children’s preferences for the normative variants?

Hypothesis 2: We hypothesise that the factors documented in the literature in different models influence children’s LAs differentially. While they are highlighted separately in different studies, combining them in a single regression model can provide a holistic and better articulated view of the emergence of LA in monolingual and bilingual children.

2. Methods

2.1. The sociolinguistic context of the study: language variation in France

France is well known for its monolingual ideology (Hélot, 2008). Significantly, the French language has a specific mention in the second article of the French constitution stating that it is the national language of France. In addition, the “standard ideology,” particularly vigorous in France, valorises uniformity in the language practices and is highly influenced by prescriptive rules, where writing norms are considered the most valued forms, associated with the production of speakers with a high social status (Gadet, 2007). The prescriptive pressures for producing correct forms as opposed to supposedly deviant ones thus influence perceptions on how speakers should speak, potentially leading to a social depreciation of alternative non-standard forms, in particular those present in spontaneous, ordinary speech. Combining a variationist approach in his analysis of the phenomena of variation in oral French, Laks (2000) discussed the sociological and structural underpinnings of different variables and illustrated how variants are placed differentially on the social scale, depending on their status in the linguistic market (Bourdieu, 1979).

Regardless of the standard ideology sketched above, sociolinguistic studies have documented variation phenomena in ordinary language practices in both adults and children. An emblematic point of variation is the negation construction, where in monitored speech both adverbial particles (*ne... pas*) are maintained whilst the *ne* deletion is characteristic of spoken French (Gadet, 1989). In their corpus study, Blanche-Benveniste and Pallaud (2001) observed that both adults and children omit the first part of the negation construction in 80% of the cases in informal interactions. Children thus bathe in heterogeneous practices where both standard and non-standard forms circulate, even in interactions in institutional contexts where the standard forms are expected (Nardy et al., 2014). This heterogeneity has been documented in previous developmental sociolinguistic studies where children at school are confronted with heterogeneous practices on a daily basis because, on the one hand, teachers tend to use more standard forms of French when adopting a teaching stance (Bigot & Maillard, 2014; Buson & Nardy, 2020) but, on the other hand, may use non-standard forms of the negation in conversations with the children or other adults present in the classroom (Buson et al., 2023). Interestingly, in Bigot and Maillard’s (2014) qualitative inquiry in primary and secondary teachers’ discourses, it was found that teachers’ production of the standard form of the negation construction increased as the learners got older. This gradual increase in the production of standard forms could suggest that teachers adjust the quantity of their monitored speech according to the learners’ stylistic and metalinguistic development.

The five linguistic constructions used in this study (see *Materials* section), documented extensively in the literature, have been chosen because (1) they all have a standard/normative and non-standard/non-normative form, and (2) they circulate in various spaces of socialisation and their availability does not depend on the exposure to a particular, geographically constrained (social) dialect.

2.2. Participants

We initially enrolled 143 monolingual and bilingual children from metropolitan France, using a snowball sampling recruitment procedure (Milroy & Gordon, 2003). Following the general personal data protection regulations (registered number 2-201108), once every parent was informed of the project, they filled out and signed a consent form

authorising the collection and the use of anonymous data. All participants were healthy individuals with no declared audition problems. Both parents' levels of education were used as a proxy for the children's SES. Children whose parents had school-level education only were assigned to the lower-class background. Children who had one parent with graduate-level education and the other parent with a lower level of education were classified as low-to-middle-class. Children whose parents had graduate-level education (bachelor's degree in the French education system) or higher were placed in the middle-to-upper-class background. This tripartite categorisation was an attempt to include a more holistic measure of the family's cultural capital through both parents' level of education, in tune with current trends where women and men are involved in childcare. Twenty-four children from the initial sample whose SES was not completed were excluded. The only child in 5th grade (153 months old) was also excluded from this initial sample. In the remaining 118 participants, the children's choices were sometimes ambiguous, as either both or neither one of the variants were selected. Ten of these ambiguous answers were produced by one 61-month-old child, who was consequently excluded from the sample. Further, all "Other" responses were excluded (the full procedure is provided in the [Supplementary Material](#)).

These answers represented 2% (N = 46) of the total number of choices (N = 2223). [Figure 1](#) gives a bird's-eye view of the distribution of the responses in our sample, according to the linguistic components under study. It clearly appears that word-final post-consonantal /s/ (PostC-R: 37/468) generated the highest number of "Other" responses.

The final sample comprised 117 children aged between 63 (5;3) and 143 (11;11) months: 46 monolingual Francophones (20 girls, 26 boys) and 71 bilingual children (35 girls and 36 boys). The mean age for girls was 103.01 months (range: 63–141, SD = 20.33) and for boys 103.93 months (range: 63–143, SD = 22.45). Approximately half of the

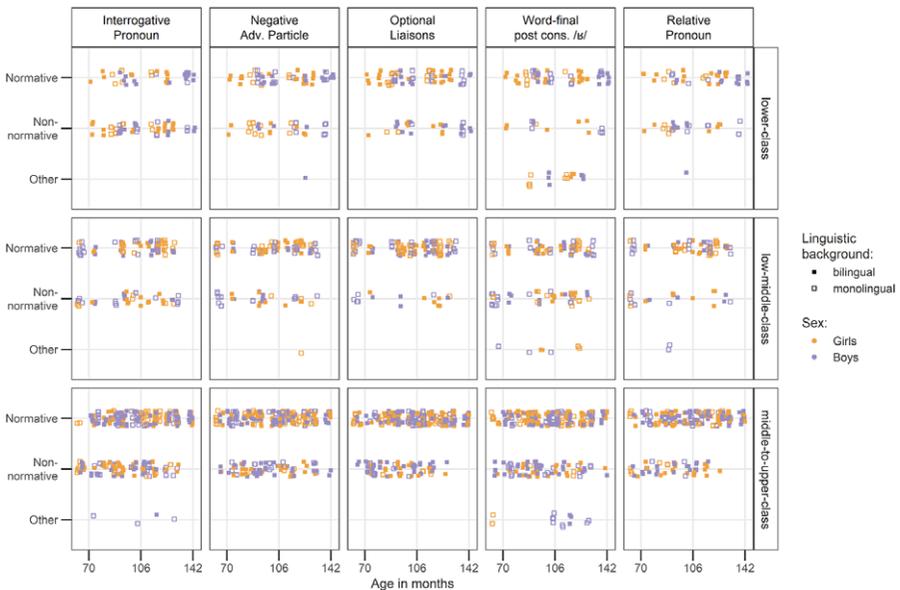


Figure 1. Response distribution in the sample (117 children, mean age = 103.5, SD = 21.39, range: 63–143).

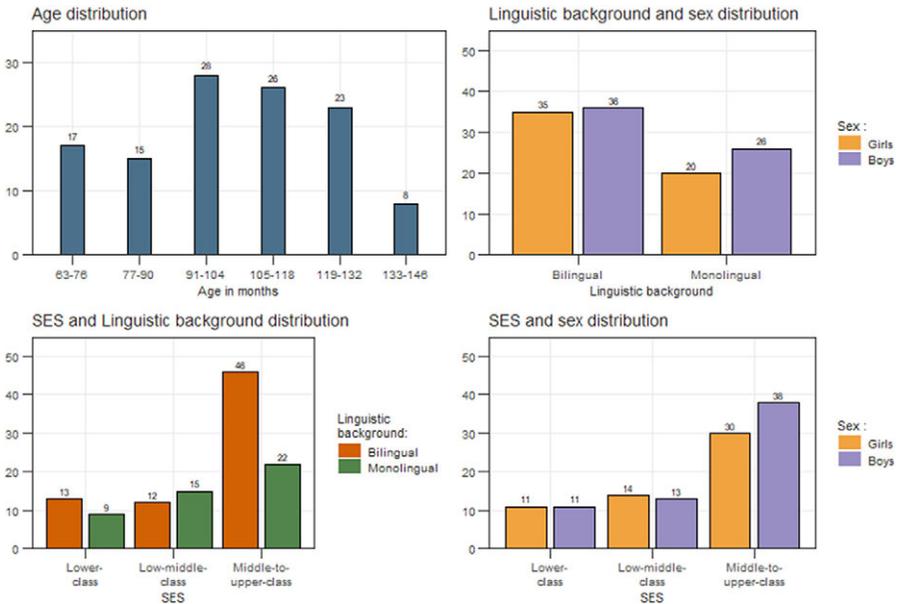


Figure 2. Distribution of the participants ($N = 117$) in the final sample according to different age groups, sex, linguistic background and SES.

cohort came from a middle-to-upper-class background. Additional information on the final sample is provided in the [Supplementary Material](#).

Figure 2 reports the distribution of the participants according to the target demographic, socioeconomic and sociolinguistic constructions of our study: age, sex, linguistic background and SES.

2.3. Bilingual participant recruitment and sample

Before presenting the bilingual children of our sample, we should consider how the monolingual/bilingual categories were operationalised in this study. First of all, it is challenging to draw a clear divide between monolingualism and bilingualism (Mackey, 1962; Surrain & Luk, 2019) or between monodialectalism and bidialectalism (Chevrot & Ghimenton, 2019). In addition, when dealing with primary school-going children's repertoires, it is equally difficult to find children who are completely monolingual, i.e., children who have had no contact with another language besides their first language. Indeed, learning of a foreign language is compulsory in the French school curriculum, and this begins at the age of 6 years. Despite these challenges, all children who spoke only French at home and attended a French-medium school were considered monolingual. Although it is a multifaceted concept comprising multiple features (Byers-Heinlein et al., 2019; Surrain & Luk, 2019), bilingualism is defined here quite categorically as the regular use of two languages in an individual's daily life (Grosjean, 2013). The bilingual children in the sample were recruited following De Houwer's (2009) classification of the patterns of bilingual acquisition and socialisation, as they were declared by their caregivers: (1) Children in the "Bilingual First Language Acquisition" (BFLA) category had two

languages from birth; (2) Children in the “Early Second Language Acquisition” (ESLA) category started with one language spoken at home and were introduced to the second language within their second birthday; (3) Children in the “Late Second Language Acquisition” (LSLA) category started with one language and then were introduced to the second language from 6 years onwards.

The 71 bilingual children in our sample had the following language acquisition profiles: 18 BFLA children; 52 ESLA children and 1 LSLA child. The ESLA category had two different languages at the onset of acquisition: it was either French or another family language to which the children were exposed first, before being introduced to a second language between ages 2 and 6 upon entering kindergarten. Three children in our sample were emergent bilinguals, as they came from Francophone families who had enrolled the child in a bilingual French–English kindergarten from age 2 onwards. The only Late Second Language child started learning French at school at age 6 and had at least 3 years of schooling in the French system. Because he attended a monolingual school, used both Lithuanian and French at home and had received at least 8 hours of French daily in a formal school setting for 3 consecutive years, we considered that he had sufficient cumulative experience of French for this particular perception study. Table 1 summarises

Table 1. Bilingual children’s language background and exposure: Home language(s) and school settings and children’s language acquisition profile (Bilingual First Language Acquisition, Early Second Language Acquisition, Late Second Language Acquisition)

BFLA (Total = 18 children)			
	<i>Bilingual School^a</i>		<i>Monolingual French School</i>
Home language(s)	French–Chinese	3	French–Chinese 15
ESLA (Acquisition of French between 2 and 6 years) (Total = 26 children)			
	<i>Bilingual School</i>		<i>Monolingual French School</i>
Home language(s)	Chinese	5	Diankhanké 3
	Spanish	1	Chinese 9
			Russian 2
			Turkish 6
ESLA (Acquisition of French from birth, L2 between 2 and 6 years) (Total = 26)			
L2 introduced after French	<i>Bilingual School</i>		<i>Monolingual French School</i>
	Chinese	2	Chinese 5
	English ^b	4	Arabic 10
			Kabyle 4
			Turkish 1
LSLA (French acquired in school at age 6 for at least 3 years) (N = 1)			
Home language(s)			<i>Monolingual French School</i>
			Lithuanian–French 1

^aThe children attending a bilingual school were exposed to both French and their home language. For example, in the first BFLA case, the three children attending a bilingual school were exposed to French and to six additional hours in Chinese in their school curriculum.

^bThree were emergent bilingual children who belonged to Francophone families and who attended a bilingual French–English school. The fourth child in this context had bilingual parents who used French at birth and then added English when the child went to a bilingual kindergarten.

the children's home language(s) and school setting. We have reported the home languages as they were declared by the children's caregivers. For example, none of the Chinese families specified the dialectal variety they spoke at home.

The children had different language pairs and came from language backgrounds that do not usually fall under the typical label of "elite bilinguals" (De Mejía, 2002), except for the 4 English–French children and one Spanish–French bilingual child. However, the question of elite bilingualism is a complex one, as elitism is socially constructed and evolves along with the host society's perceptions and ideologies. For example, Codó and Sunyol (2018) documented the changes in discourses of elitism produced in a Spanish community towards a Chinese community, through an ethnography of children enrolled in an international school in Barcelona. Despite the complexity underpinning elite bilingualism, this was not considered to be problematic, as our study was not centred on language proficiency and maintenance of the minority family language but was focused rather on the perception and categorisation of variation in French, the language that all the children had in common and to which they had exposure in school and in wider social circles.

In addition, in the French context, the majority of schools have a monolingual language instruction mode, yet some schools have a bilingual program where children receive six additional hours of language instruction in another language (see Ghimenton et al. (2023) for a study on the bilingual French–English habitus in children enrolled in a French international primary school). The children in our sample were enrolled in a bilingual program that involved their home language, except for the three emergent bilinguals who came from Francophone families but were exposed to English since they were 2 years old.

2.4. Materials

A PowerPoint presentation was used to project the experiment using a laptop computer. Each slide contained two identical female faces on a white background (see Figure 1 in Supplementary Material). A female vocal stimulus was associated with each female face. Each slide corresponded to one trial in which two sentences (stimuli) were presented, one after the other. Both expressed the same semantic meaning, but one was a normative realisation of one of the five phonetic or syntactic constructions, and the other a non-normative realisation. The constructions are well documented in both sociolinguistic and psycholinguistic literature, see Table 2 (see also Müller (1989) and Gadet (2007) for a comprehensive account of variation in French). For example, there were two realisations of the sentence *il est très embêté*: in the first possible realization, the optional liaison was realised [ilɛtʁɛzãbɛtɛ], whereas in the second it was not [ilɛtʁɛãbɛtɛ]. The stimuli (see Supplementary Material for the stimuli list) were produced by the same female adult talker who was a French L1 speaker and was familiar with both linguistic variants. Each sentence lasted no longer than 3 seconds. The stimuli were presented in counterbalanced order (slide order and normative/non-normative presentation order). Five linguistic constructions were considered: their abbreviated forms appearing in Table 2 will be henceforth used in the body of this paper. The terms "component"/"construction" are used rather than "variable" for the sake of clarity, since in our statistical models, NEG, IP, RP, PostC-R and OL are the five levels of a single categorical variable playing the role of a predictor of the participants' choices.

Table 2. The five linguistic components investigated in the study

Abbreviated form/ component	Examples of variation N = Normative; NN = Non-normative	References
NEG/Negative adverbial particle	Je n'ai pas fini. (N) J'ø ai pas fini. (NN)	Gadet (2007)
IP/Interrogative Pronoun Qui/C'est qui qui	Qui a mangé la brioche ? (N) C'est qui qui a mangé la brioche ? (NN)	Buson (2008)
RP/Relative pronoun Dont/Que	C'est le livre dont je te parle. (N) C'est le livre que je te parle. (NN)	Buson (2008)
PostC-R/Word-final post-consonantal /ɤ/ /ɤ/	Il y en a quatre. (N) [iljānakatɤ] Il y en a quat(re). (NN) [iljānakat]	Chevrot et al. (2000)
OL/Optional liaison	Il est très ([z]) embêté. (N) [ilɛtɤzɛzübɛtɛ] Il est très ø embêté. (NN) [ilɛtɤzɛübɛtɛ]	Barbu et al. (2013); Nardy (2008); Nardy et al. (2014)

As mentioned in the sociolinguistic presentation of the context, the linguistic variables included in this set are all typical of the French metropolitan area and circulate in various spaces of socialisation, including school, as the alternation between normative and non-normative forms is the fruit of inter- and intra-individual stylistic variation.

2.5. Procedure

Children first performed a working memory task involving numbers and words in order to test whether they were capable of storing two sentence stimuli in their working memory (BALE – *Batterie analytique du langage écrit*, www.cognisciences.com). All the children were capable of performing the test, and once completed, they participated in the experiment comprising 20 trials (4 trials per linguistic construction). Before the experiment started, the investigator explained the task to the child and ran two practice trials. Following Barbu et al. (2013) experimental design, the task was introduced by the investigator as follows: “In each slide, you will hear two sentences, one after the other. After hearing both sentences, you must say which one of the two speakers spoke best.” For each trial, the investigator showed the pair of identical female faces and played the voice stimuli by pointing at the corresponding face. Once the child heard both stimuli, s/he pointed towards the stimulus that best matched the answer to the investigator’s question. The task was filmed in order to capture the child’s choice, that either corresponded to a normative or non-normative production. Once the child completed all 20 trials, the investigator asked what s/he thought of the sentences. For one of the variables, the RP *dont/que*, only 3 trials were taken into consideration due to a technical failure. Our final results are thus based on 19 sentence pairs and not on 20, as initially expected.

2.6. Modelling speaker preferences

Adopting a confirmatory approach to assess predefined hypotheses, we performed statistical modelling of the participants’ preferences toward the normative or non-normative variants of the 5 linguistic components.

We worked with R (R Development Core Team, 2023) and the following packages of functions: *lme4* (Bates et al., 2015), *lmerTest* (Kuznetsova et al., 2017), *DHARMa* (Hartig, 2022), *emmeans* (Lenth, 2023), *ggplot2* (Wickham, 2016), and *sjPlot* (Lüdtke, 2023). The workflow used to analyse this sample is available in the [Supplementary Material](#).

We considered a mixed-effects logistic regression model (Barth & Kapatsinski, 2018; Levshina, 2015; Speelman, 2014) to account for the participants' two possible answers: preference for the normative or the non-normative variant of the linguistic component presented. We defined our binary dependent variable, RESPONSE, as follows: 0 means preferring the non-normative variant and 1 preferring the normative variant.

We included the following fixed predictors, which translate both our linguistic cases of interest and the several demographic, socioeconomic, and sociolinguistic factors we want to investigate:

- The categorical variable LING_VAR, with 5 levels corresponding to the different linguistic components under study: NEG, IP, RP, PostC-R, and OL
- The categorical variable SES, with 3 levels: lower-class, low-middle-class, and middle-to-upper-class
- The categorical variable SEX, with two levels: boy and girl
- The categorical variable MONO_BI, with two levels: monolingual and bilingual
- The continuous variable AGE in months

One can note that many studies in LA treat age as a categorical variable (after binning), but beyond, in many sociolinguistic and psycholinguistic studies, this choice is by and large unmotivated (Bowie, 2014; Eckert, 2014; Macaulay, 2009). We chose to keep age as continuous since binning means losing potentially valuable information and scaled it in the model to ease convergence.

Given our research questions and our willingness to conduct an integrative investigation of our various parameters, in addition to main effects we also included in our model all possible 10 (simple/double) interactions between the predictors: LING_VAR:SES, LING_VAR:SEX, LING_VAR:MONO_BI, LING_VAR:AGE, SES:SEX, SES:MONO_BI, SES:AGE, SEX:MONO_BI, SEX:AGE, and MONO_BI:AGE. We did not consider higher-level interactions due to the complexity of their analysis.

Based on the literature reported above, we expected:

- a facilitator effect of SES: the higher the class, the more adherence to the normative variants
- an interaction between AGE and SEX: girls prefer normative forms before boys
- an interaction between AGE and MONO_BI: bilingual children prefer the normative forms before monolinguals

With respect to the non-independence of our observations given the by-subject and by-item (i.e., in our case, by-sentence pairs) groupings, we considered random effects in addition to the previous fixed effects. Our key variables were named ID for the participants (with 117 levels) and SENTENCE_PAIR for the pairs of sentences (with 19 levels). To remain conservative when formulating conclusions, we compared different models to determine the best random-effects structure – with appropriate random intercepts and random slopes – given our fixed effects, considering Barr (2013) and Barr et al. (2013)'s maximalist stance and Bates et al. (2018)'s more pragmatic recommendations.

We eventually ended with a converging model with two crossed random intercepts for ID and SENTENCE_PAIR and a single categorical random slope: LING_VAR for ID. The step-by-step procedure is described in detail in [Supplementary Material](#).

Following the usual syntax in *lme4* in R, our final model can be written as:

```
RESPONSE ~ LING_VAR + scale(AGE) + MONO_BI + SEX + SES +
LING_VAR:SES + LING_VAR:SEX + LING_VAR:MONO_BI + LING_VAR:AGE +
SES:SEX + SES:MONO_BI + SES:AGE + SEX:MONO_BI + SEX:AGE + MONO_BI:AGE +
(0 + LING_VAR | ID) + (1 | SENTENCE_PAIR)
[Family: binomial (logit)]
```

We verified that our model satisfied the assumptions of logistic regressions. We then assessed the significance of the various fixed effects. In particular, we derived estimated marginal means for the groups corresponding to the various levels of our categorical predictors in order to assess significant differences between them. *p*-Values were adjusted when performing multiple tests.

3. Results

We first investigated the higher-level predictors of the model, i.e., the ten two-way interactions. The first four interactions tested were between one categorical variable and the continuous variable (age). Given that none of these interactions were significant, we focused on age as a main effect. It was highly significant ($z = 4.407, p < .0001$), with children gradually growing into preferring the normative variants, all other things considered equal. A one-month increase in age was associated with a 0.0216 increase on the log-odds linear scale (95% confidence interval [0,012–0,0312]). [Figure 3](#) reports predicted values on the probability scale for the different levels of the categorical variables. This figure also illustrates that the probability of preferring the normative variant for OL among monolingual girls was very high regardless of age. Conversely, the increase in preference for the normative variants over time was most obvious for RP. One can further note the children's indecision concerning IP, particularly among girls and boys from the lower and low-middle-class.

As for interactions between categorical variables, we first found the effect of the linguistic background on children's preference for the normative variants to be modulated by sex. Multiple comparisons with a "mvt"¹ adjustment suggested that a single difference was significant, between monolingual and bilingual girls ($z = -2.784, p = .0273$) (see [Figure 4](#)). For monolingual girls, the probability of choosing the normative variants reached 83.5% (95% CI [73.5%–90.2%]), while it was only 69.9% (95% CI [59.9%–78.3%]) for bilingual girls (averaging the effect of the linguistic variable, age, and SES).

Regarding the 2×3 interaction between linguistic background and SES, two significant pairwise differences involved monolinguals from the middle-to-upper class. As illustrated in [Figure 5](#), their probability of preferring the normative variants (85.5%, 95% CI [75.9%–91.3%]) was significantly higher than that of bilinguals (65.1%, 95% CI [48.1%–79.1%]) ($z = -3.251, p = .014$) and monolinguals from the low-middle-class (62.7%, 95% CI

¹According to the manual for the *emmeans* package, the 'mvt' adjustment for multiple comparisons is 'the closest to being the "exact" all-around method "single-step" method, as it uses the multivariate *t* distribution (and the *mvtnorm* package) with the same covariance structure as the estimates to determine the adjustment.' (<https://cran.r-project.org/web/packages/emmeans/vignettes/confidence-intervals.html>).

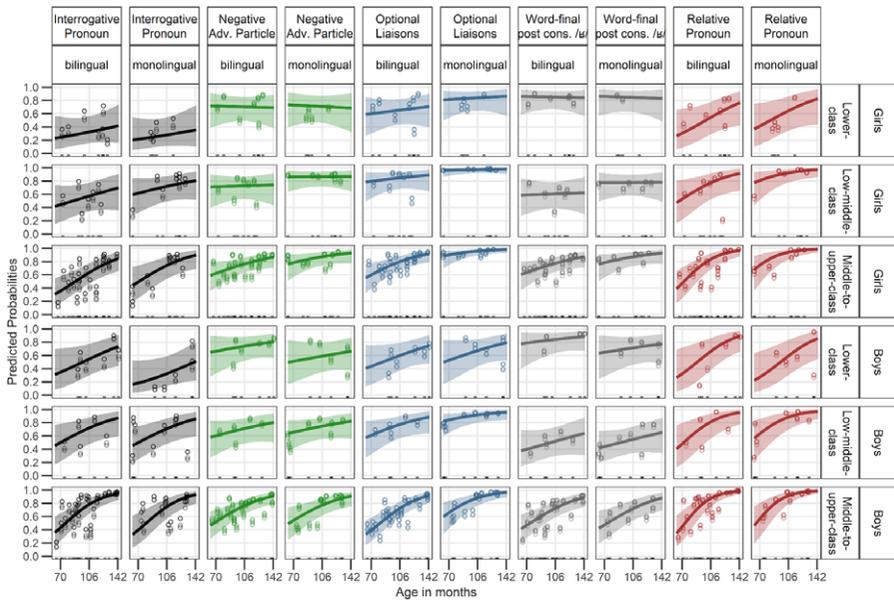


Figure 3. Individual predicted probabilities of preference for the normative variants and population mean curves by age, linguistic variable, sex, linguistic background and SES. Each mean curve reported is based on fixed effects only, i.e., all random effects are set to zero. Ninety-five percent confidence intervals (CIs) are also based on the fixed effects and their uncertainty only.

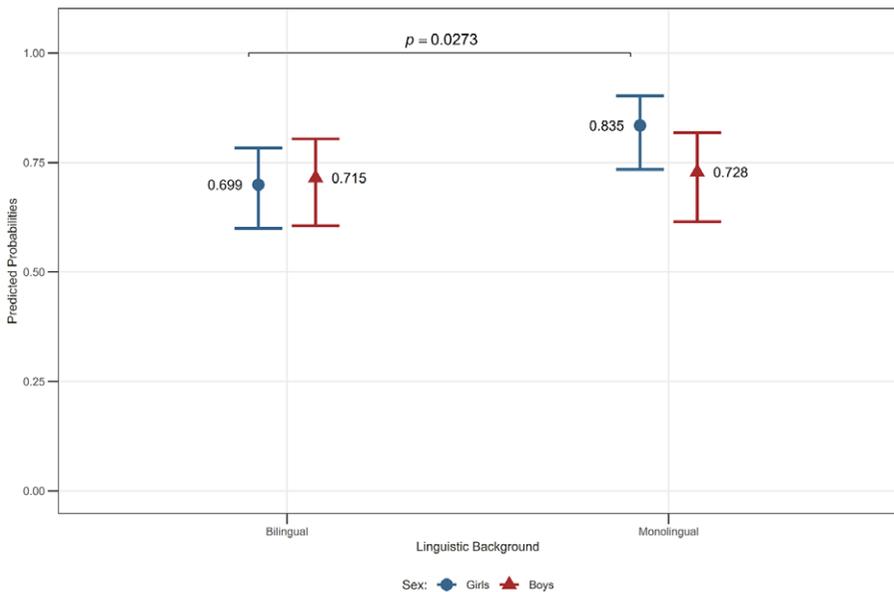


Figure 4. Linguistic Background \times Sex Interaction plot for predicted probabilities of favouring the normative variants, with 95% confidence error bars. The single significant contrast is indicated at the top of the figure.

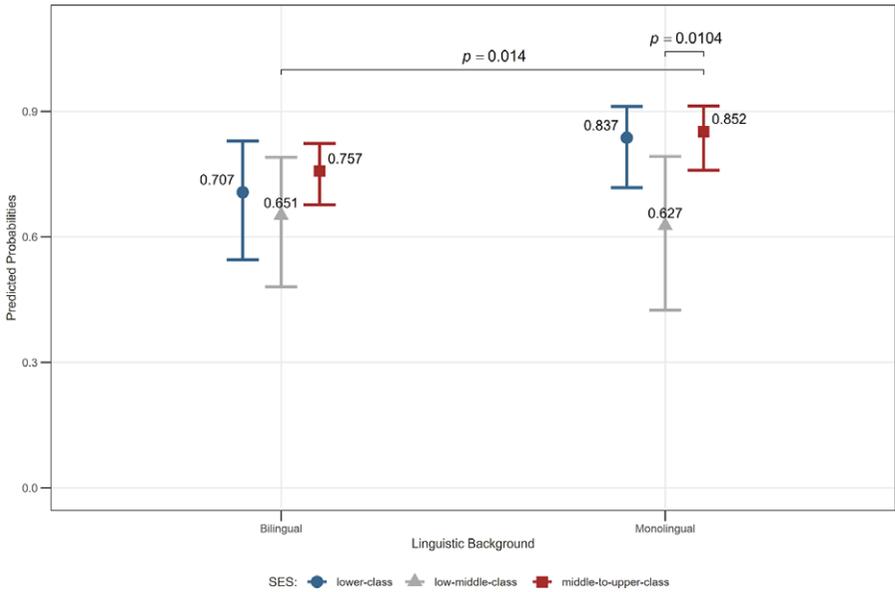


Figure 5. Linguistic Background \times SES Interaction plot for predicted probabilities of favouring the normative variants, with 95% confidence error bars. Only significant contrasts are indicated.

[42.5%–79.3%]) ($z = -3.336, p = .0104$). Interestingly, only the 95% confidence interval of the marginal mean for low-middle-class monolingual and bilingual children included 50%, and thus one cannot reject the hypothesis of a possible random choice for these groups.

As for the 2×3 interaction between sex and SES (see Figure 6), the probability of preferring the normative variants was significantly higher for girls from the middle-to-upper class (83.2%, 95% CI [74.3%–89.5%]) than for girls (65.4%, 95% CI [47.0%–80.2%]) ($z = -2.854, p = .0475$) and boys (62.4%, 95% CI [42.5%–78.8%]) ($z = -3.042, p = .0275$) from the low-middle-class. No significant differences emerged between the other social conditions. The 95% confidence interval of the marginal means for low-middle-class girls and boys both include 50%.

Children’s preferences towards the normative variants were also modulated by the 2×5 interaction between linguistic backgrounds and linguistic variables (averaging the effect of sex, SES, and age). As illustrated in Figure 7, of all the five linguistic variables investigated, the probabilities of preferring the normative variant were the lowest for IP for both bilinguals and monolinguals: 58.7% (95% CI [40.2%–75.1%]) among bilinguals 57.9% (95% CI [37.1%–76.3%]) among monolinguals. These two 95% CIs included 50%. Conversely, the probability of preferring the normative variant was the highest for OL among monolinguals (90.1%, CI [78.2%–95.8%]). The pairwise tests indicated that the predicted preference for the normative variants was significantly higher for monolinguals than for bilinguals (73.2%, CI [57.7%–84.5%]) for OL ($z = -3.409, p = .0216$). The former was also significantly higher than the probability of favouring the normative variant for the IP in bilinguals ($z = -4.307, p = .0007$) and in monolinguals ($z = -4.078, p = .0019$).

Turning to the 2×5 interaction between sex and linguistic variables, as illustrated in Figure 8, the probability of preferring the normative variant was the highest for OL in girls (88%, 95% CI [75.6%–94.5%]). This probability was significantly higher than that of

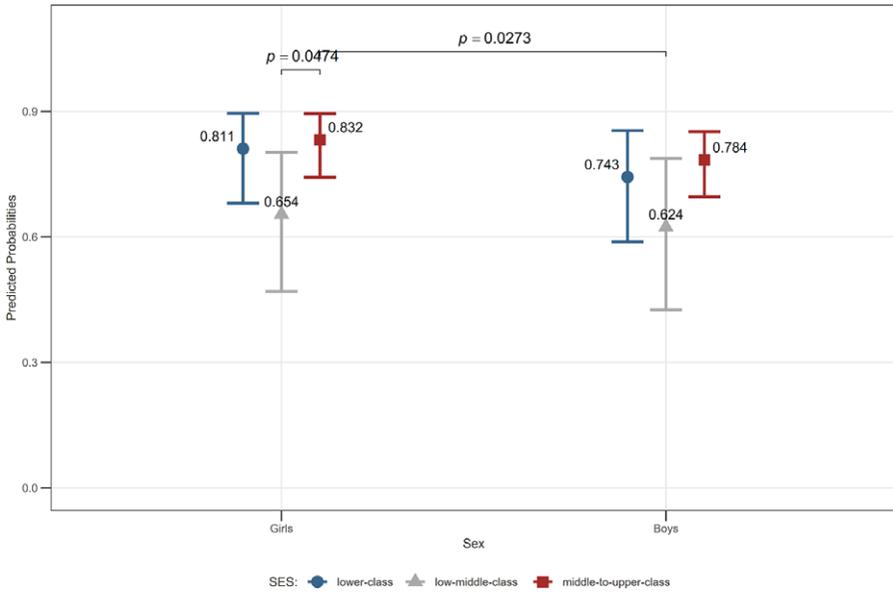


Figure 6. Sex x SES Interaction plot for predicted probabilities of favouring the normative variants, with 95% confidence error bars. Only significant contrasts are indicated.

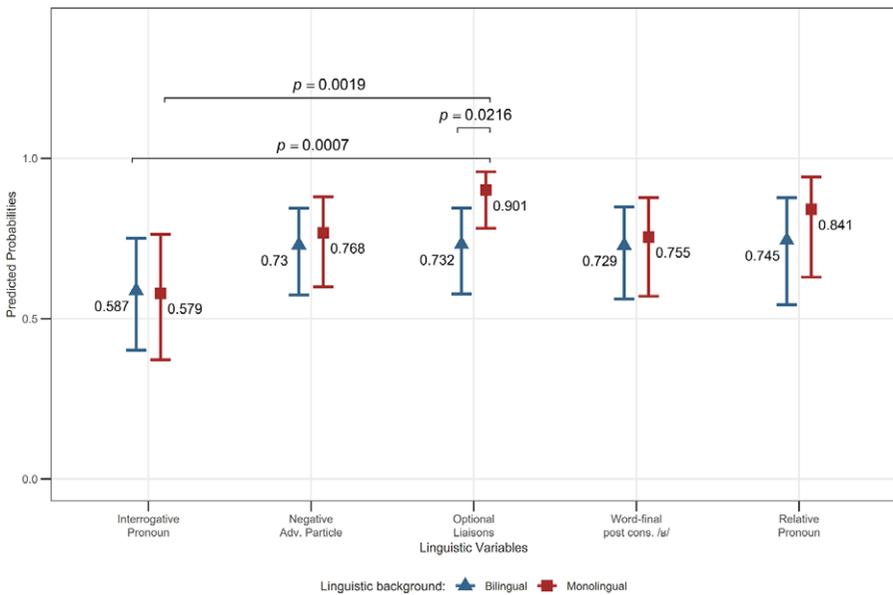


Figure 7. Linguistic variable x Linguistic background Interaction plot for predicted probabilities of favouring the normative variants, with 95% confidence error bars. Only significant contrasts are indicated.

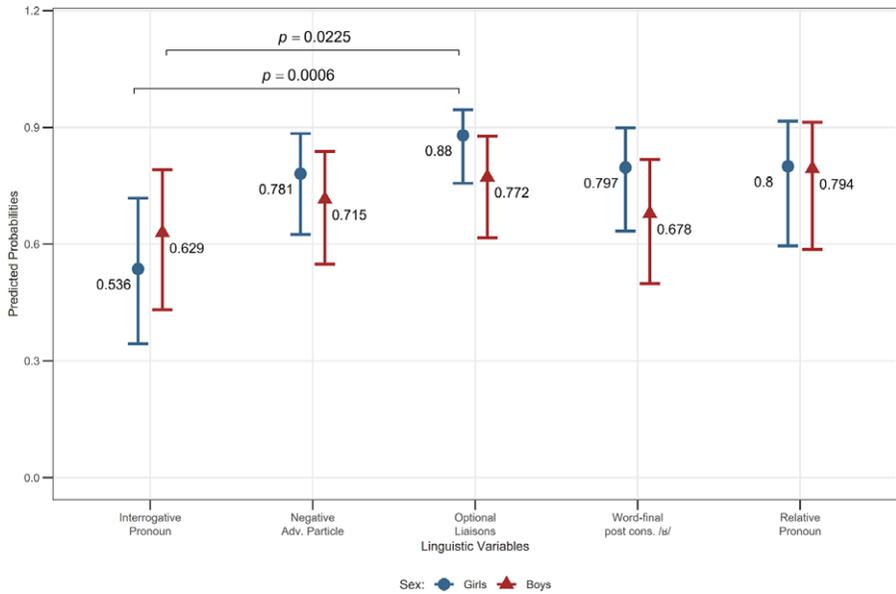


Figure 8. Linguistic variable \times Sex Interaction plot for predicted probabilities of favouring the normative variants, with 95% confidence error bars. Only significant contrasts are indicated.

preferring the normative variant for IP in both girls ($z = -4.349$, $p = .0006$) and boys ($z = 3.397$, $p = .0225$). It is also worth mentioning that, as with the previous two-way interaction between linguistic variables and linguistic backgrounds, the probabilities of preferring the normative variant were lowest for IP, both in girls (53.6%, 95% CI [34.4%–71.8%]) and boys (62.9%, 95% CI [43.2%–79.1%]). The 95% confidence interval for these two marginal means comprised, once again, 50%.

The 3×5 interaction between SES and linguistic variables was more difficult to interpret. The first feature to point out is that in the intermediate social class, the probability of preferring the normative variant for the IP was the lowest (35.4%, 95% CI [14.4%–64.2%]) and this induced eight of the nine significant contrasts found in total (see Figure 9).

Within the intermediate social group, with the exception of the PostC-R level, for which the probability of preferring the normative variant was 81.7% (95% CI [55.7%–94.1%]), the CIs for the marginal means of the other four levels included 50%. Conversely, among children from middle-to-upper social classes, the CIs for the five linguistic variables did not include 50%, suggesting that within 95% of cases, children from this social group preferred the normative variants. Within the lower class, only the CIs for the PostC-R and IP levels included 50%. More generally, it is noteworthy that for the IP, NEG, OL and RP levels, the preferences of lower-class children were very close to those of middle-to-upper-class children. Conversely, the behaviour of low-middle-class children showed a different, and above all random, response pattern, since the CIs for each of these four levels included zero. In more peripheral terms, the contrast between the probability of preferring the normative variant for OL and for PostC-R was also significant in the lower social class ($z = 3.382$, $p = .0469$).

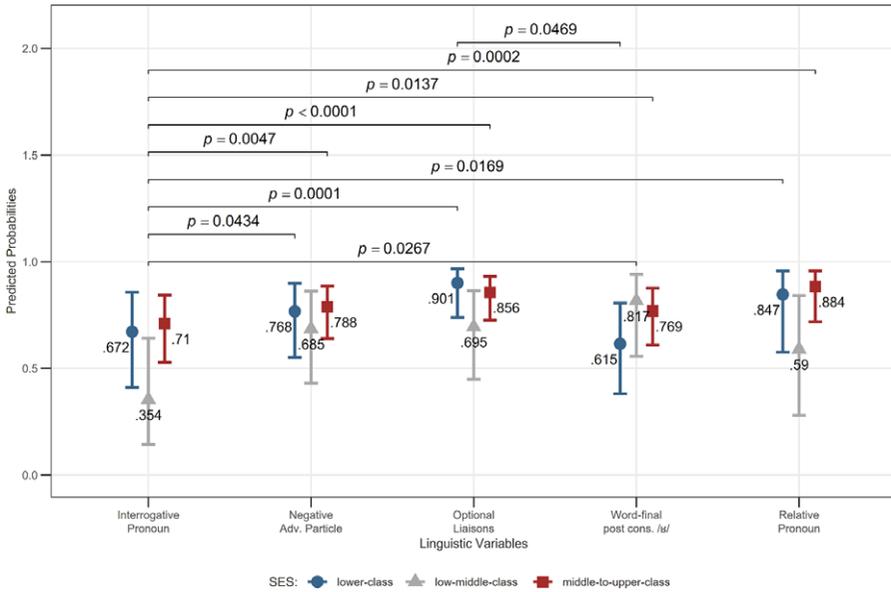


Figure 9. Linguistic variable \times SES Interaction plot for predicted probabilities of favouring the normative variants, with 95% confidence error bars. Only significant contrasts are indicated.

4. Discussion

Summing up very briefly the general results, our mixed-effects model revealed several statistical differences. Age was the main effect; thus, it was possible to conclude that children’s preferences for normative variants increased with age. Although no overall significant differences appeared between boys and girls, the interactions found between factors revealed a more complex picture. Preferences for the normative variants were significantly higher for monolingual girls than for bilingual girls and also higher for girls from the middle-to-upper class than for both low-to-middle-class girls and boys.

These results shed light on our two research questions, enabling us to assess the two respective hypotheses.

Concerning our first research question inquiring on the way children’s LAs unfold over time, we found that age was a highly significant predictor explaining the evolution of LAs, enabling the conclusion that normative variants are acquired over time through the children’s language socialisation experiences. However, the emerging preferences for the normative variants had different developmental slopes depending on the linguistic construction. This means that each linguistic construction documented was subject to a different preference timeframe. When investigating each linguistic construction in detail, it is worth noting that the differences in the probability of preferring the normative variant for OL and IP were such that they induced many significant contrasts. When averaging across all the other variables, only the 95% confidence interval of the marginal mean of the IP level included 50% (58.3%, 95% CI [44.1%–71.3%]). Hence, the hypothesis of a random choice concerning the IP construction cannot be rejected. All in all, while this study corroborated the fact that, with age, children start preferring the normative variants, the pivotal age of 8 documented in the literature was not confirmed for all the linguistic components. The first hypothesis was thus confirmed: children’s LAs vary depending on

the linguistic construction evaluated, just as each linguistic level (phonetic, morphological, syntactic, etc.) is subject to different acquisition patterns. In this case, the probabilities of preferring the normative variant for OL (morphophonological level) were the highest from the onset (63 months), whilst the level of indecision was the highest for the IP construction (syntactic level) throughout the age range (63–141 months).

Moving on to the second research question, we wanted to know which factors explained the possible variation patterns observed in the children's preferences for the normative variants. Because in the developmental literature, different factors have been shown to influence the patterns in language production acquisition patterns, we hypothesised that these same factors could also influence our participants' LA trajectories. Importantly, instead of documenting the factors in different models, we considered them all within the same model so that we could obtain a more holistic yet robust and explicative view of the emergence of LA in monolingual and bilingual children. For the sake of clarity, we will briefly discuss each factor, excluding age, separately, and then we will move on to the general discussion.

4.1. Sex

No overall significant differences appeared between boys and girls, and this may be surprising in the light of the results in the literature. Indeed, previous studies on production dealing with sociolinguistic variation in language acquisition (Barbu *et al.*, 2014) have shown that boys produced more vernacular variables than girls and that boys were exposed to more non-normative variants than girls (Foulkes *et al.*, 2005). However, the interactions found between sex, and respectively, the children's linguistic background and their family's SES shed light on the complexity of LA mechanisms. For example, the probabilities of preferring the normative variants were particularly high for monolingual girls from privileged social backgrounds. We discuss this point further below.

4.2. Linguistic background

Regarding the differences noted between bilinguals and monolinguals, our results are in contrast with other studies where bilinguals were found to develop preferences for the prestigious varieties before monolinguals (Day, 1980; Rosenthal, 1974). However, in these studies, children were asked to express their preferences towards two distinct varieties, whereas in our study, children had to evaluate variants *within* one language variety. Bilinguals may require more time to develop clear-cut preferences towards the normative variants in the French context, where both are regularly used in everyday conversations. In general, given their wider exposure to language variation in the input (intra- and interlinguistic), Francophone bilinguals may manifest preferences towards the French normative (prescriptive) variety at a later stage compared to monolinguals who have a narrower set of variants in competition in their environment (intralinguistic only). These variants, together with their indexical values, would become entrenched earlier than other, less frequent ones (Langacker, 1988).

4.3. SES

Previous research conducted with the same type of preference-based task with Franco-phone monolingual children showed that only the children from a high-SES background

preferred the normative variants (Barbu et al., 2013; Nardy, 2008). Our sample comprised slightly older children with respect to the aforementioned studies, and we found slightly different patterns. For the results involving SES, lower-class and upper-class children had similar preference patterns for most of the linguistic components under study. Yet, low-to-middle-class children's preferences showed the most contrasts and random responses. These results support the tripartite coding system we used for SES, as it may help shed light on another sociolinguistic phenomenon that has seldom been considered in the attitudinal literature: The lower-to-middle-class children's indecision regarding their preferences for the normative variants may illustrate what Labov (1972) called linguistic insecurity. Indeed, children in this category had parents who came from two different SES, backgrounds and this may have exposed them to language practices and representations anchored in different social allegiances. In other words, these children belonged to families that were likely to have different social class legitimacies which may have contributed to a general lack of confidence towards the assignment of their preferences on the basis of what form was considered to be socially most desirable.

4.4. General discussion: Reading the attitudes through the lens of language socialisation via exposure and experience

Considering our set of results, schooling and exposure to the written medium are undoubtedly important factors pushing children's preferences towards the normative variants. The negative adverbial particle (NEG) illustrated this point: The sociolinguistic literature showed that the absence of the first element of negation is highly frequent in spoken French, irrespective of the situation's level of formality (Gadet, 2007; Müller, 1989). Hence, because children start with exposure to the spoken language, their familiarity with the non-normative variant is substantially higher than their exposure to the normative variant. Once they go to school, however, their increasing exposure to the written medium raises awareness of the first element of variation *ne* (which is compulsory in the written medium). Exposure to the written norms that contrast with the norms underpinning the forms to which they are most accustomed could result in competing preferences towards two different levels of legitimacy during the first stages of schooling.

Contrastively, it is possible to interpret the results obtained for the LAs towards OL by considering the effects of frequency of exposure. Although children are exposed to fluctuations between uses of the normative and the non-normative realisations, unlike for NEG, they are exposed to these fluctuations very early in their lives. Hence, they have more time to build socio-cognitive representations of the different uses of the variable. Said otherwise, earlier exposure to the normative and non-normative variants could familiarise the children with the societal preferences towards the more prestigious variants during the early stages of language socialisation. Consequently, this would lead to the earlier emergence of language preferences towards the more prestigious variant. The indecision noted in the LAs towards the IP component is the most intriguing. As a syntactic component, LAs towards IP would require more language exposure before they are stabilised. However, unlike with NEG, the other syntactic component, there may be less discourse produced on the IP constructions because these are not linked to a formal and compulsory rule in writing. There could also be a complementary explanation for the particular status of the IP component. Bonhomme (2021) explored the different IP constructions found in six CHILDES (MacWhinney, 2000) corpora: Champaud (1994); Leveillé (Suppes et al., 1973); MTLN (Le Normand et al., 2013); Palasis (2009);

Pauline (Bassano *et al.*, 2004); York/Anne (De Cat & Plunkett, 2002). The corpora considered comprised spontaneous interactions between children (1–4-year-olds) and their caregivers. Although the age groups did not correspond to those in our sample, they nonetheless provided an estimate of the typical Francophone child's cumulative exposure to, and experience of, the interrogative pronoun use prior to the ages tested in our experiment. It was found that in spoken language produced in ordinary conversations, there were different constructions expressing interrogation: (1) *qui*+verb realisations (normative); (2) *c'est qui qui*; (3) *qui c'est qui* or (4) *qui est-ce qui*. It could be argued then that, faced with these variable forms that circulate in different spaces of socialisation, children take longer to “prefer” one as opposed to another variant because the normative and non-normative forms are not in a strict binary relation with one another (normative variant versus non-normative variant) and have other competitors.

In more general terms, LAs would be the fruit of the accumulation of language socialisation experiences via the input but also of the exposure to the societal ideologies that influence the way forms are interpreted as more desirable than others. Because ideologies circulate within the spaces of socialisation to which children participate and where they are exposed, these ideologies are reproduced in discourses that shape and reinforce children's social representations and forms of prescription in a rather top-down fashion (e.g. with the acquisition of writing, children are systematically exposed to discourse on the compulsory nature of the first particle of the NEG construction). Children's input influences the elaboration of their sociolinguistic representations (bottom-up process), and these representations consequently shape the children's perceptions and attitudes. Such bottom-up processes are driven by factors such as frequency of exposure to a particular use of one or more variants of linguistic constructions. As for the top-down processes, they also encompass language ideologies regarding language prestige where one particular variant, usually the normative variant, is preferred over another, perceived as less prestigious (Garrett, 2010; Milroy, 2001). The language ideologies and representations carried in child-directed discourse or in normative discourses produced at school preempt the non-normative variants produced by the child or in the child's environment, reinforcing ideals of correctness. Hence, this would explain why it is likely that the child is socialised into preferring variants that do not necessarily match the ones found in her production, nor those in her social group. Both top-down and bottom-up processes depend on time, frequency of exposure and on the indexical values a social group establishes between linguistic features and a speaker's characteristics. Through these bottom-up and top-down influences (Fuchs & Hofkirchner, 2005), the sociocognitive representations of each variable become progressively stabilised and attuned with the social pressures for the promotion of the prestigious variants, contributing to the children's socialisation process.

4.5. Conclusion: Limitations and future directions

One of the study's limitations, which is currently being addressed, is an investigation where the collection of perception data is complemented with production data in order to better comprehend how the linguistic constructions circulate in monolingual and bilingual children's spaces of socialisation. Future directions should consider combining self-declared and objective language proficiency measures and dominance within the methodological set-up. This would shed light on the potential linguistic insecurity effects hypothesised in this study and paint a finer-grained picture of the development of LA.

All in all, the interactions emerging in our statistical model highlight the complexity of the construct “language attitudes towards variation.” Our holistic approach is a first attempt to account for this complexity through a single model explaining the factors influencing the developmental trajectory of social cognition.

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/S0305000925000169>.

Acknowledgements. The authors are grateful to the ASLAN project (ANR-10-LABX-0081) of the Université de Lyon for its financial support within the French program “Investments for the Future” operated by the National Research Agency (ANR). The first author created the experimental design, supervised the data collection, and wrote the discussion. The second and the corresponding authors created the statistical models and performed the statistical analyses. The corresponding author also curated the data, interpreted the results and prepared all the figures. The third and fourth authors collected part of the data and prepared the databases from which the analyses were performed.

Competing interests. The authors declare none.

Disclosure of use of AI tools. An AI tool (OpenAI ChatGPT 4o mini) was used to check that the formatting of the bibliography met APA standards.

References

- Anisfeld, E., & Lambert, W. E. (1964). Evaluational reactions of bilingual and monolingual children to spoken languages. *The Journal of Abnormal and Social Psychology*, *69*(1), 89. <https://doi.org/10.1037/h0040913>
- Barbu, S., Martin, N., & Chevrot, J.-P. (2014). The maintenance of regional dialects: A matter of gender? Boys, but not girls, use local varieties in relation to their friends’ nativeness and local identity. *Frontiers in Psychology*, *5*, 1215. <https://doi.org/10.3389/fpsyg.2014.01251>
- Barbu, S., Nardy, A., Chevrot, J.-P., & Juhel, J. (2013). Language evaluation and use during early childhood: Adhesion to social norms or integration of environmental regularities? *Linguistics*, *51*(2), 381–411. <https://doi.org/10.1515/ling-2013-0015>
- Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*. <https://doi.org/10.3389/fpsyg.2013.00328>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Barth, D., & Kapatsinski, V. (2018). Evaluating logistic mixed-effects models of corpus-linguistic data in light of lexical diffusion. In D. Spielman, K. Heylen, & D. Geeraerts (Eds.), *Mixed-effects regression models in linguistics: Quantitative methods in the humanities and social sciences* (pp. 99–116). Springer. https://doi.org/10.1007/978-3-319-69830-4_6
- Bassano, D., Laaha, S., Maillochon, I., & Dressler, W. U. (2004). Early acquisition of verb grammar and lexical development: Evidence from periphrastic constructions in French and Austrian German. *First Language*, *24*(1), 33–70. <https://doi.org/10.1177/0142723704040704>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models (v2). *arXiv:1506.04967 [stat]*. <http://arxiv.org/abs/1506.04967>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bigot, V., & Maillard, N. (2014). “Dites pas c’que j’dis, dites c’que j’écris...”: Représentations et pratiques d’enseignants vis-à-vis de la variation en contexte scolaire. *Lidil*, *50*, 81–104. <https://doi.org/10.4000/lidil.3577>.
- Blanche-Benveniste, C., & Pallaud, B. (2001). Le recueil d’énoncés d’enfants : Enregistrements et transcription. *Recherches sur le français parlé*, *16*, 11–37.

- Bonhomme, N.** (2021). *Le développement des attitudes linguistiques chez l'enfant vis-à-vis de la variation. Etude exploratoire autour de trois variables syntaxiques du français : 'dont', 'ne', 'qui'*. [Master thesis, Université Lumière Lyon 2 (Lyon)]
- Bouchard Ryan, E., & Giles, H.** (Eds.). (1982). *Attitudes towards language variation*. Hodder Arnold.
- Bourdieu, P.** (1979). *La distinction*. Minuit.
- Bowie, D.** (2014). Age: Methods and metadata. *Language and Linguistics Compass*, 8(11), 519–528. <https://doi.org/10.1111/lnc3.12121>
- Buson, L.** (2008). La variation stylistique chez les enfants de 10/11 ans: Une étude exploratoire en contexte français. In M. Abecassis, L. Ayosso, & E. Vialleton (Eds.), *Le français parlé au XXIème siècle (Normes et variations géographiques et sociales)* (Vol. 1, pp. 203–222). Paris: L'Harmattan.
- Buson, L., & Nardy, A.** (2020). Le(s) français scolaire(s) à l'école maternelle: Pratiques langagières d'enseignant.e.s et d'élèves. *Le français aujourd'hui*, 208, 93–103.
- Buson, L., Nardy, A., Muller, D., & Chevrot, J.-P.** (2018). The sociolinguistic repetition task: A new paradigm for exploring the cognitive coherence of language varieties. *Topics in Cognitive Science*, 10(3), 803–817. <https://doi.org/10.1111/tops.12380>
- Buson, L., Nardy, A., Rousset, I., & Zhang, C.** (2023). Style in a school community – “Ne” deletion in French preschool. *Journal of Sociolinguistics*, 27(4), 384–401. <https://doi.org/10.1111/josl.12611>
- Byers-Heinlein, K., Behrend, D. A., Said, L. M., Girgis, H., & Poulin-Dubois, D.** (2017). Monolingual and bilingual children's social preferences for monolingual and bilingual speakers. *Developmental Science*, 20(4), e12392. <https://doi.org/10.1111/desc.12392>
- Byers-Heinlein, K., Esposito, A. G., Winsler, A., Marian, V., Castro, D. C., & Luk, G.** (2019). The case for measuring and reporting bilingualism in developmental research. *Collabra: Psychology*, 5(1), 37. <https://doi.org/10.1525/collabra.183>
- Champaud, C.** (1994). The development of verb forms in French children at around two years of age: Some comparisons with Romance and non-Romance languages. In *First lisbon meeting on child language*, Lisbon, Portugal, (pp. 14–17).
- Chervel, A.** (1977). *Et il fallut apprendre à écrire à tous les petits Français: Histoire de la grammaire scolaire*. Payot.
- Chervel, A.** (2008). *L'orthographe en crise à l'école: Et si l'histoire montrait le chemin?* Retz.
- Chevrot, J.-P., Beaud, L., & Varga, R.** (2000). Developmental data on a French sociolinguistic variable: The word-final post-consonantal /R/. *Language Variation and Change*, 12(3), 295–319. <https://doi.org/10.1017/S095439450012304X>
- Chevrot, J.-P., & Ghimenton, A.** (2019). Bilingualism and bidialectalism. In A. De Houwer & L. Ortega (Eds.), *The Cambridge handbook of bilingualism* (pp. 510–523). Cambridge University Press.
- Codó, E., & Sunyol, A.** (2018). 'A plus for our students': The construction of Mandarin Chinese as an elite language in an international school in Barcelona. *Journal of Multilingual and Multicultural Development*, 40(5), 436–452. <https://doi.org/10.1080/01434632.2018.1543694>
- Coronel-Molina, S. M.** (2014). Definitions and critical literature review of language attitude, choice, and shift; Samples of language attitude surveys. *Journal of Second and Multiple Language Acquisition*, 2(2), 25–77.
- Cremona, C., & Bates, E.** (1977). The development of attitudes toward dialect in Italian children. *Journal of Psycholinguistic Research*, 6(3), 223–232. <https://doi.org/10.1007/BF01068021>
- Day, R. R.** (1980). The development of linguistic attitudes and preferences. *TESOL Quarterly*, 14(1), 27–37. <https://doi.org/10.2307/3586806>
- De Cat, C., & Plunkett, B.** (2002). Qu'est-ce qu'i (l) dit, celui-là? Notes méthodologiques sur la transcription d'un corpus francophone. In *Romanistische Korpuslinguistik: Korpora und gesprochene Sprache/Romance Corpus Linguistics: Corpora and Spoken Language (CD-rom)*. Tübingen: Narr.
- De Houwer, A.** (2009). *Bilingual first language acquisition*. Multilingual Matters.
- De Mejía, A. M.** (2002). *Power, prestige, and bilingualism: International perspectives on elite bilingual education*. Multilingual Matters. <https://doi.org/10.21832/9781853595929>
- DeJesus, J. M., Hwang, H. G., Dautel, J. B., & Kinzler, K. D.** (2017). Bilingual children's social preferences hinge on accent. *Journal of Experimental Child Psychology*, 164, 178–191. <https://doi.org/10.1016/j.jecp.2017.07.005>

- Dekker, S. V., Duarte, J., & Loerts, H. (2021). 'Who really speaks like that?' – Children's implicit and explicit attitudes towards multilingual speakers of Dutch. *International Journal of Multilingualism*, **18**(4), 551–569. <https://doi.org/10.1080/14790718.2021.1908297>
- Eagly, A. H., & Chaiken, S. (1998). Attitude structure and function. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vols. 1–2, 4th ed., pp. 269–322). McGraw-Hill.
- Ekert, P. (2014). The problem with binaries: Coding for gender and sexuality. *Language and Linguistics Compass*, **8**(11), 529–535. <https://doi.org/10.1111/lnc3.12113>
- Fehér, K. (2020). Status-based preference of varieties in bidialectal kindergarteners: An experimental study. *Argumentum*, **16**, 147–173. <https://doi.org/10.34103/ARGUMENTUM/2020/10>
- Foucault, M. (1975). *Surveiller et punir*. Gallimard.
- Foulkes, P., Docherty, G. J., & Watt, D. (2005). Phonological variation in child-directed speech. *Language*, **81**(1), 177–206. <https://doi.org/10.1353/lan.2005.001>
- Fuchs, C., & Hofkirchner, W. (2005). The dialectic of bottom-up- and Top-down-emergence in social systems. *tripleC*, **3**(2), 28–50. <https://doi.org/10.31269/triplec.v3i2.21>
- Gadet, F. (1989). *Le français ordinaire*. Armand Colin.
- Gadet, F. (2007). *La variation sociale en français* (rev. and expanded ed.). Ophrys.
- Garrett, P. (2010). *Attitudes to language* (1st ed.). Cambridge University Press.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, **8**(1), 8–11. <https://doi.org/10.1016/j.tics.2003.10.016>
- Ghimenton, A., Chevrot, J.-P., & Billiez, J. (2013). Language choice adjustments in child production during dyadic and multiparty interactions: A quantitative approach to multilingual interactions. *Linguistics*, **51**(2), 413–438. <https://doi.org/10.1515/ling-2013-0016>
- Ghimenton, A., Cohen, C., & Minniear, J. N. (2023). Constructing a bilingual French-English habitus through language experience: A socialisation account from children attending a bilingual programme at a state school in France. *Journal of Multilingual and Multicultural Development*, 1–17. <https://doi.org/10.1080/01434632.2023.2218353>
- Giles, H., Hewstone, M., & Ball, P. (1983). Language attitudes in multilingual settings: Prologue with priorities. *Journal of Multilingual and Multicultural Development*, **4**(2–3), 81–100. <https://doi.org/10.1080/01434632.1983.9994104>
- Girard, F., Floccia, C., & Goslin, J. (2008). Perception and awareness of accents in young children. *British Journal of Developmental Psychology*, **26**(3), 409–433. <https://doi.org/10.1348/026151007X251712>
- Grosjean, F. (2013). Bilingualism: A short introduction. In F. Grosjean & P. Li (Eds.), *The psycholinguistics of bilingualism* (pp. 5–28). John Wiley & Sons.
- Häcki Buhofer, A., & Studer, T. (1993). Zur Entwicklung von Sprachdifferenzbewusstsein und Einstellung zu den Varianten des Deutschen in der Deutschen Schweiz. *Bulletin CILA (Schweizerische Hochschulkommission für angewandte Sprachwissenschaft) (Bulletin VALS-ASLA seit 1994)*, **58**, 179–199.
- Hartig, F. (2022). *DHARMA: Residual diagnostics for hierarchical (multi-level/mixed) regression models*. R package version 0.4.6. <https://CRAN.R-project.org/package=DHARMA>
- Hélot, C. (2008). Bilingual education in France: School policies versus home practices. In C. Hélot & A.-M. de Meija (Eds.), *Forging multilingual spaces: Integrated perspectives on majority and minority bilingual education* (pp. 203–227). Multilingual Matters.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, **26**(1), 55–88. <https://doi.org/10.1016/j.dr.2005.11.002>
- Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In M. H. Bornstein (Ed.), *Handbook of parenting* (2nd ed., pp. 231–251). Erlbaum.
- Kaiser, I., & Kasberger, G. (2014). Children's sociolinguistic preferences: The acquisition of language attitudes within the Austrian standard-dialect-continuum. In *ViLA2 conference, Grenoble, 3–5 December 2014*.
- Kaiser, I., & Kasberger, G. (2018). Children's emerging ability to discriminate L1-varieties. *First Language*, **38**(5), 447–480. <https://doi.org/10.1177/0142723718761403>
- Kinzler, K. D., & DeJesus, J. M. (2013). Northern = smart and Southern = nice: The development of accent attitudes in the United States. *Quarterly Journal of Experimental Psychology*, **66**(6), 1146–1158. <https://doi.org/10.1080/17470218.2012.731695>

- Kinzler, K. D., Shutts, K., & Spelke, E. S. (2012). Language-based social preferences among children in South Africa. *Language Learning and Development: The Official Journal of the Society for Language Development*, 8(3), 215–232. <https://doi.org/10.1080/15475441.2011.583611>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Labov, W. (1972). *Sociolinguistic patterns*. Blackwell.
- Lafontaine, D. (1986). *Le parti pris des mots*. Mardaga.
- Laks, B. (2000). De la variation et des variantes: À propos du relâchement. *Linx* [online], 42. <https://doi.org/10.4000/linx.747>
- Lambert, W. E., Frankle, H., & Tucker, G. R. (1966). Judging personality through speech: A French-Canadian example. *Journal of Communication*, 16(4), 305–321. <https://doi.org/10.1111/j.1460-2466.1966.tb00044.x>
- Langacker, R. W. (1988). A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics* (pp. 127–161). Benjamins. <https://doi.org/10.1075/cilt.50.06lan>
- Le Normand, M.-T., Moreno-Torres, I., Parisse, C., & Dellatolas, G. (2013). How do children acquire early grammar and build multiword utterances? A corpus study of French children aged 2 to 4. *Child Development*, 84(2), 647–661. <https://doi.org/10.1111/j.1467-8624.2012.01873.x>
- Lenth, R. (2023). *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.8.8. <https://CRAN.R-project.org/package=emmeans>
- Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. Benjamins.
- Lüdtke, D. (2023). *sjPlot: Data visualization for statistics in social science*. R package version 2.8.15. <https://CRAN.R-project.org/package=sjPlot>
- Macaulay, R. (2009). *Quantitative methods in sociolinguistics*. Palgrave Macmillan.
- Mackey, W. T. (1962). The description of bilingualism. *Canadian Journal of Linguistics*, 7, 51–85.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Lawrence Erlbaum Associates.
- Milroy, J. (2001). Language ideologies and the consequences of standardization. *Journal of Sociolinguistics*, 5(4), 530–555. <https://doi.org/10.1111/1467-9481.00163>
- Milroy, L., & Gordon, M. (2003). *Sociolinguistics: Method and interpretation*. Blackwell. <https://doi.org/10.1002/9780470758359.ch1>
- Müller, B. (1989). *Le Français d'aujourd'hui* (1re éd.). Klincksieck.
- Nardy, A. (2008). *Acquisition des variables sociolinguistiques entre 2 et 6 ans : Facteurs sociologiques et influences des interactions au sein du réseau social* [Doctoral dissertation, Université Stendhal (Grenoble)]. <http://www.theses.fr/2008GRE39056>
- Nardy, A., Chevrot, J.-P., & Barbu, S. (2014). Sociolinguistic convergence and social interactions within a group of preschoolers: A longitudinal study. *Language Variation and Change*, 26(3), 273–301. <https://doi.org/10.1017/S0954394514000131>
- Ochs, E. (1990). Indexicality and socialization. In J. W. Stigler, R. A. Shweder, & G. Herdt (Eds.), *Cultural psychology: Essays on comparative human development* (pp. 287–308). Cambridge University Press. <https://doi.org/10.1017/CBO9781139173728.009>
- Palasis, K. (2009). *Syntaxe générative et acquisition: Le sujet dans le développement du système linguistique du jeune enfant* [Doctoral dissertation, Université de Nice (France)].
- Preston, D. (1998). They really speak bad English down South and in New York City. In L. Bauer & P. Trudgill (Eds.), *Language myths* (pp. 139–149). Penguin Books.
- R Development Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <http://www.r-project.org>
- Rosenthal, M. S. (1974). The magic boxes: Preschool children's attitudes toward Black and Standard English. *Florida FL Reporter*, 55, 93.
- Rosseeel, L., Speelman, D., & Geeraerts, D. (2018). Measuring language attitudes using the Personalized Implicit Association Test: A case study on regional varieties of Dutch in Belgium. *Journal of Linguistic Geography*, 6(1), 20–39. <https://doi.org/10.1017/jlg.2018.3>
- Sim, J. H., & Post, B. (2024). Early phonological acquisition in multi-accent contexts. In E. Babatsouli (Ed.), *Multilingual acquisition and learning: Towards an ecosystemic view to diversity* (pp. 194–216). Benjamins.
- Souza, A. L., Byers-Heinlein, K., & Poulin-Dubois, D. (2013). Bilingual and monolingual children prefer native-accented speakers. *Frontiers in Psychology*, 4, 953. <https://doi.org/10.3389/fpsyg.2013.00953>

- Speelman, D.** (2014). Logistic regression: A confirmatory technique for comparisons in corpus linguistics. In D. Glynn & J. A. Robinson (Eds.), *Human cognitive processing* (Vol. 43, pp. 487–533). Benjamins. <https://doi.org/10.1075/hcp.43.18spe>
- Suppes, P., Smith, R., & Leveillé, M.** (1973). The French syntax of a child's noun phrases. *Archives de Psychologie*, 42, 207–269.
- Surraïn, S., & Luk, G.** (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005 and 2015. *Bilingualism: Language and Cognition*, 22(2), 401–415. <https://doi.org/10.1017/S1366728917000682>
- Vaughn, C., & Becker, K.** (2024). Documenting the emerging social-semiotic landscape in children ages 5 to 12. *Language & Communication*, 95, 16–30. <https://doi.org/10.1016/j.langcom.2024.01.001>
- Wickham, H.** (2016). *ggplot2: Elegant graphics for data analysis* (2nd ed.). Springer. <https://doi.org/10.1007/978-3-319-24277-4>.

Cite this article: Ghimenton, A., Coupé, C., Bonhomme, N., Song, J., & Arnaud, V. (2025). Modeling monolingual and bilingual children's language attitudes towards variation in metropolitan France. *Journal of Child Language* 1–27, <https://doi.org/10.1017/S0305000925000169>