Prediction in bilingual sentence processing

Is it linked to production?

Natalia Meir, Olga Parshina, and Irina A. Sekerina
Bar-Ilan University | Middlebury College | The City University of New York

The *Unified Competition Model* (MacWhinney, 2012) accounts for cross-linguistic differences in thematic role mapping. We investigated production and predictive use of accusative case morphology in Russian-Hebrew bilingual children. We also investigated the role of production in predictive processing testing the *Prediction-by-Production Account* (Pickering & Garrod, 2018) vs. the *Missing Surface Inflection Hypothesis* (Prévost & White, 2000). Three groups of children aged 4–8 participated: Russian-Hebrew-speaking bilinguals, Russian-speaking and Hebrew-speaking monolingual controls. All children participated in the accusative case production and Visual-World eye-tracking comprehension experiments. Bilinguals were tested in both of their languages. The results of the study confirmed the predictions of the *Unified Competition Model* showing typological differences in the strength of the case-marking cue and its predictive use in sentence processing in Russian- and Hebrew-speaking controls. While Russian-speaking monolinguals relied on case marking to predict the upcoming agent/patient, the performance of Hebrew-speaking monolingual children varied. The findings for bilinguals showed that despite their lower production accuracy in both languages, they were either indistinguishable from monolinguals or showed an advantage in the predictive use of case morphology. The findings support the *Missing Surface Inflection Hypothesis*, which predicts a dissociation between production and comprehension.

Keywords: sentence processing, prediction, Russian, Hebrew, case morphology, Visual-World, Visual-World paradigm (eye-tracking)

https://doi.org/10.1075/lab.22102.mei
Linguistic Approaches to Bilingualism ISSN 1879-9264 | E-ISSN 1879-9272

1. Introduction[1]

There are substantial well-documented systematic cross-linguistic differences in the mapping of thematic roles, such as the AGENT of the action and the THEME/PATIENT of the action (e.g., *Who did what to whom*?), onto the arguments. Speakers of different languages rely on different cues (e.g., word order, animacy, subject-verb agreement, and case markings) to interpret simple sentences with transitive verbs (e.g., *The rabbit sees a fox.*) (MacWhinney et al., 1984). In such sentences, the subject is typically an animate AGENT, and the object is an inanimate PATIENT (Allen, 2014). In the languages that do not use morphological case marking (e.g., English), the most reliable cue to this mapping is a word-order cue, and the noun-verb-noun strings are strongly biased towards a Subject-Verb-Object (SVO) interpretation. However, in the languages with flexible word orders (e.g., Turkish and German), speakers rely on case-marking cues, thus enabling Object-Verb-Subject (OVS) interpretations (MacWhinney & Bates, 1989).

According to *the Unified Competition Model* (MacWhinney, 2012), language-specific cues for thematic-role assignment differ in availability (how frequently does a speaker encounter that particular cue in the input?), reliability (does the cue always signal the same relation?), processing cost (how difficult is that cue?), and conflict resolution (can the cue be overridden?). The model explains different processing strategies regarding conflict resolution between word order and morphological case marking in monolingual adults. However, processing strategies differ depending on the cue weight in a specific language which determines the speed of processing: strong cues lead to faster responses and competing cues lead to slower ones. In OVS sentences, the agent is sentence-final, which creates a conflict between canonical subject-first word order and the case marking, resulting in the need to assign different weights to the conflicting cues.

With respect to children, there is no agreement on how they resolve such cue conflicts and to what extent their processing strategies are adult-like. Although previous studies showed that children are less efficient than adults in using morphosyntactic cues, it is well-established that this ability improves with age. For Russian monolingual children aged 4–6, previous off-line studies demonstrated that they assign more weight to the case-morphology cue than word order when processing OVS sentences (Janssen & Meir, 2019; Sauermann & Gagarina, 2018).

1. A preliminary iteration of the study was published in the BUCLD proceedings (see Meir et al., 2020). The current version introduces a more extended scope, incorporating an additional subset of research questions and hypotheses. Furthermore, the dataset of the current study included three groups of children all tested using identical procedures and eye-tracking apparatus.

Moreover, Aumeistere et al. (2022) in an eye-tracking study found that 3-to-6-year-olds used the grammatical gender cue in an anticipatory manner in real time. In contrast, German-speaking 5-year-old children rely less on case marking and require additional cues to correctly comprehend sentences with the non-canonical word order, while 7-year-olds can resolve conflicting cues in an adult-like manner (Dittmar et al., 2008). Similarly, earlier off-line studies in Hebrew also showed that word order is stronger than case marking; children only gradually learn to refrain from relying solely on word order and start paying attention to the accusative (ACC) (ACC) case marker *et*, gender, and number agreement: differences between 3-, 4-, 5-, 6-, 7-, and 10- year-olds were reported (e.g., Frankel et al., 1980; Frankel & Arbel, 1981).

1.1 Predictive processing in children: Comprehension vs. production

Prediction is one of the key mechanisms of fluent language processing (e.g., Ferreira & Chantavarin, 2018; Kuperberg & Jaeger, 2016). It occurs when a comprehender activates upcoming linguistic information (e.g., anticipates an argument) before s/he encounters it in the input. According to Pickering and Gambi (2018), the underlying mechanism to anticipate upcoming information in comprehension relies on the same mechanisms which are used in sentence production, as per the *Prediction-by-Production Account* (Martin et al., 2018; Pickering & Garrod, 2007, 2013). A comprehender relies on his/her own production system which triggers the retrieval and generates production representations, constituting the comprehender's prediction of the speaker's upcoming utterance.

There is direct and indirect evidence in favor and against the *Prediction-by-Production Account* (Pickering & Gambi, 2018). On the one hand, Brouwer et al. (2017) showed that monolingual Dutch children who were more accurate in producing correct gender-marked articles were also able to predict the upcoming noun on its basis, thus confirming the link between production and comprehension. In bilingual children, there is indirect evidence for the link between production and comprehension: bilingual children err more than monolinguals in case-marking production (Meir et al., 2021; Meir & Janssen, 2021; Schwartz & Minkov, 2014), as well as show reduced case-marking comprehension (Chondrogianni & Schwartz, 2020; Gagarina & Klassert, 2018; Janssen & Meir, 2019). On the other hand, there is evidence for a dissociation between production and comprehension (Grimm et al., 2011; Hendriks & Koster, 2010) that supports the *Missing Surface Inflection Hypothesis* (Prévost & White, 2000). For example, bilinguals can be sensitive to ungrammaticalities on par with monolinguals while still making considerable production errors (Blom et al., 2016; Chondrogianni & Marinis, 2012).

On-line studies in the Visual World eye-tracking Paradigm (hereafter VWP; see Kamide et al., 2003; Sekerina, 2014, for more details) allow us to investigate how cues are integrated in real-time, whether they are used in conflict resolution and in prediction. VWP studies that test the predictive use of morphosyntactic cues in monolinguals, i.e., whether case morphology is integrated as the utterance unfolds, show that monolingual children rely on morphosyntactic information to predict the upcoming linguistic elements before they are encountered in speech. For example, Özge et al. (2019) reported that Turkish-speaking adults and 4-year-old children used the case-marking cue in verb-medial and -final sentences to predict the upcoming noun although children were slower. They extended this finding to German-speaking children aged 4;0–4;06 (Özge et al., 2022) and showed that they too were sensitive to the ACC case morphology and interpreted sentence-initial case-marking cues independent of the verb and the word order (SOV and OSV) in predicting the thematic role of the upcoming argument.

In the current study, we investigated the relationship between production and comprehension in predictive on-line processing of simple transitive sentences in monolingual and bilingual children whose two languages, Russian and Hebrew, differ in morphological cue weight. Russian provides multiple case cues, e.g., case agreement between nouns and their modifiers, simultaneously allowing for considerable flexibility in word order. Unlike Russian, in Hebrew, the ACC case marking is limited to definite contexts only, and there is no case agreement between the noun and its modifiers. Thus, this difference in cue weight between these languages provides a testing ground for comparison of monolingual Russian- and Hebrew-speaking children that can shed light on cross-linguistic differences in monolingual predictive processing. The comparison of Russian-Hebrew-speaking bilingual children to their monolingual peers will advance our understanding of bilingual sentence processing, and how cues in the two languages of bilinguals might interact. In the following subsection, we briefly overview ACC case realization in Russian and Hebrew.

1.2 Accusative case marking and word order in Russian and Hebrew

Russian realizes case features that differ morphologically for three grammatical genders and three noun declension classes. Table 1 lists Nominative (NOM) and Accusative (ACC) inflections across the 1st and the 2nd declensions. The 3rd declension class is omitted because it was not included in the current study. The current study employed animate and inanimate nouns of all three genders. On some nouns (e.g., FEM nouns of the 1st declension and MASC animate nouns of the 2nd declension), a dedicated inflection is used for the ACC case, which is different from the NOM case, thus this form is overtly marked. However, on other nouns, the ACC case inflection is homophonous to the NOM case (e.g., inanimate nouns of the 2nd declension, both masculine and neuter), thus on these forms the ACC case is unmarked. All modifiers, including adjectives, numerals, and demonstratives, agree with the noun in case and carry their own case-marking morphemes (Shvedova, 1980).

Table 1. The Russian case inflections ([NOM] → [ACC]) for the nouns of 1st and 2nd declension classes

Declension	Features	NOM	ACC	English translation
1st Declension	FEM/ +Animate	*lis-a*	*lis-**u***	fox
	FEM/ -Animate	*morkovk-a*	*morkovk-**u***	carrot
	MASC/ +Animate	*pap-a*	*pap-**u***	daddy
2nd Declension	MASC/ +Animate	*zajčik*	*zajčik-**a***	bunny
	MASC/ -Animate	*mjod*	*mjod*	honey
	NEUT/-Animate	*okn-o*	*okn-o*	window

Case marking is critical for word-order variation that is characteristic of Russian grammar. Typically, Russian subjects are marked with the NOM case, whereas direct objects are marked with the ACC case (Shvedova, 1980). In a simple transitive sentence, Russian permits all six basic word orders (Bailyn, 2012). SVO is the default form and statistically the most frequent: SVO 63.3%; OVS 18.2%; OSV 14.4%; SOV 2.5%; VOS 1%; VSO 0.6% (Slioussar & Makarchuk, 2022). Case marking is the strongest cue in reversible transitive sentences for monolingual adult Russian and bilingual heritage Russian speakers (Kempe and MacWhinney, 1998; Ionin et al., 2023, accordingly). Only 5% of transitive sentences in Russian are ambiguous, e.g., when both subject and direct object are inanimate and masculine (e.g., *Tractor pot͡sarapal avtobus* '(the) tractor.NOM/ACC scatched (the) bus.NOM/ACC') (see Kempe and MacWhinney, 1998). Because case morphology in these rare sentences cannot be used to disambiguate thematic roles, the word-order cue is prioritized, i.e., SVO is preferred over OVS.

Monolingual Russian-speaking children show high accuracy in case production of familiar nouns already by age 3, although the acquisition of some irregular forms might continue up to age 6–7 (Cejtlin, 2009). Bilingual children who acquire Russian in contact with another language, such as English, German, Hebrew, and Dutch, lag behind (Chrabaszcz et al., 2023; Gagarina & Klassert, 2018; Meir & Janssen, 2021; Schwartz & Minkov, 2014). Janssen and Meir (2019) conducted an off-line study that investigated children's sensitivity to the ACC case cue in reversible transitive sentences. When word-order and case cues competed (e.g., OVS sentences), monolingual Russian children correctly comprehended OVS sentences by prioritizing case over word order. In contrast, bilingual Russian-Hebrew and Russian-Dutch children were less accurate in using case-marking cues.

Hebrew does not use case inflections on nouns but marks the ACC case by the particle *et* before definite nouns, as illustrated in (1a–b) for different contexts (Danon, 2001; Hacohen et al., 2021). For indefinite nouns, modern Hebrew has a strict SVO word order; however, OVS and OSV word orders (2) are possible with definite subjects and objects, but they are secondary, i.e., derived by syntactic movement (Friedmann & Shapiro, 2003).

(1) a. SVO – Indefinite Context:
ha- arnav ro'e šual
DEF_rabbit sees.M.SG.3P fox
'The rabbit sees a fox.'

b. SVO – Definite Context:
ha- arnav ro'e et ha-šual
DEF-rabbit sees.M.SG.3P ACC DEF-fox
'The rabbit sees the fox.'

(2) a. OVS – Definite Context:
et ha-šual ro'e ha-arnav
ACC DEF-fox sees.M.SG.3P DEF_rabbit
'The rabbit sees a fox.'

Previous studies demonstrated that in production, monolingual Hebrew-speaking children do not omit the ACC case marker *et* in front of definite nouns (Berman, 1981). Using off-line comprehension tasks, Frankel et al. (1980) showed that the Hebrew ACC marker *et* is a reliable marker, yet younger children do not always regard it as a valid cue and follow the word-order strategy (Frankel & Arbel, 1981; Sokolov, 1988). Biran and Ruigendijk (2015) found that comprehension of SVO sentences in the forced-choice task by 5–6-year-old children was close to ceiling, but comprehension of OVS and OSV sentences was lower. A recent study using an acceptability judgement task showed that Hebrew-speaking 7-year-old children start demonstrating sensitivity to the ACC case, yet their judgments are still not adult-like (Plaut & Hacohen, 2022).

1.3 The current study

Mechanisms of weighing cues in prediction have recently received a lot of attention in sentence comprehension research with monolingual children (for an overview see Pickering & Gambi, 2018), yet studies on bilingual children's morphosyntactic prediction are less numerous and mainly focus on predictive use of gender cues (but see Bosch & Foppolo, 2022, 2023; Bosch et al., 2022; Lemmerth & Hopp, 2019). Therefore, more research is needed to evaluate morphosyntactic prediction in bilingual children (for a review see Karaca et al., 2021). Russian-Hebrew bilingualism offers an opportunity to test how case marking and word order that are weighted differently in the two languages (Russian, the heritage language, HL, and Hebrew, the societal language, SL) interact in bilingual children. The current study investigated production and on-line comprehension of the ACC case morphology in transitive sentences in both languages of bilinguals and compared them to monolingual children. An earlier version of the study appeared in the BUCLD proceedings (see Meir et al., 2020) with a partial overlap in participants. For the current study, we formulated the following research questions:

RQ1 Are there differences between Russian-Hebrew bilingual and Russian-speaking and Hebrew-speaking monolingual children in the production of ACC case marking?

To test our RQ1, elicited production tasks targeting ACC case in Russian and in Hebrew were administered (see 2.2.2). Based on the previous literature, we hypothesized that bilingual children would be less accurate in both languages.

RQ2 Are there differences between Russian-Hebrew bilingual and Russian-speaking and Hebrew-speaking monolingual children in on-line comprehension?

To test our RQ2, VWP eye-tracking experiments were conducted in Russian and in Hebrew (see 2.2.3) to investigate whether participants can predict the upcoming second noun phrase (NP2) in the sentence once the case morphology of the first NP (NP1) is available. RQ2 can be separated into two hypotheses: (2a) for monolingual and (2b) for bilingual children. Based on *the Unified Competition Model* (MacWhinney, 2012), we expect:

~~(2)~~ a. For monolingual Russian-speaking children, the case-marking cue will be stronger than the word-order cue. They will be accurate and fast in predicting NP2 based on the morphology (NOM/ ACC) of NP1. For monolingual Hebrew-speaking children, the *et* ACC marking will not be the strongest cue, and they will be less accurate and slower in predicting the upcoming NP2.

b. For bilinguals, we hypothesize several possibilities. The patterns of predictive looks of bilingual children will resemble those of their monolingual peers, but bilinguals may be slower. Alternatively, cues in the two languages of bilingual children will interact. Participants will either show poorer sensitivity to the ACC case in HL-Russian as a result of transfer from the dominant SL-Hebrew in which case cue is weighted lower than word order. Or, bilingual children will show enhanced sensitivity to the ACC case in Hebrew as a result of transfer from the HL-Russian cue weight.

RQ3 To what extent does age affect the prediction in comprehension in monolingual and bilingual children?

To test our RQ3, our monolingual and bilingual samples included children in the age range of 4–8 years. Monolingual children will become more accurate in prediction as they grow older (Aumeistere et al., 2022). For bilingual children, there is inconclusive evidence for the effect of age on HL and SL development (Paradis, 2023): increasing age might be associated with better performance in the SL, yet a decline and/or fossilization in the HL are also possible.

RQ4 To what extent is prediction in comprehension linked to production mechanisms in monolingual and bilingual children?

As per the *Prediction-by-Production Account* (Pickering & Gambi, 2018), if bilingual children show low accuracy in production of the ACC case, then they will not use the ACC case cue predictively in comprehension. Alternatively, if there is a dissociation between production and comprehension (Grimm et al., 2011), bilingual children may make errors in production of the ACC case, yet they will reliably use case morphology in comprehension; this will in turn support the *Missing Surface Inflection Hypothesis* (Prévost & White, 2000).

2. Method

The data that support the findings of this study are openly available in OSF at: https://osf.io/fapdm/?view_only=74c98955deef4b15824d13d3af4a7bd3

2.1 Participants

Participants included 57 children[2] aged 4–8 in three groups: Russian-Hebrew bilinguals (BILING), monolingual Russian-speaking (MonoRU) and Hebrew-speaking controls (MonoHE). We opted for a slightly larger age range in order to capture differences and/or similarities in developmental trajectories in monolingual and bilingual children. BILING and MonoHE children were recruited in Israel. All BILING participants were born in Israel to Russian-speaking families and acquired Russian as their HL and Hebrew as their SL. MonoRu controls were also tested in Israel; they were from Russian-speaking families that immigrated to Israel from the Russian Federation. They lived in Israel between 0–3 months and had no command of Hebrew. All children were typically developing, as determined by parental questionnaires, and had no history of developmental disorders (language impairment, autism spectrum disorder, etc.). All had normal or corrected-to-normal vision, and none of the children had any history of hearing impairment.

Demographic data and language background measures were collected via a shortened version of the *Bilingual Parent Questionnaire (BIPAQ)* (Abutbul-Oz & Armon-Lotem, 2022), see Table 2. The questionnaire provides detailed demographic information (e.g., age, gender, parental education, country of birth, number of siblings in the family) as well as information on previous and current HL-SL linguistic experience and languages spoken by the child). Parents of monolingual and bilingual children filled out the same questionnaire subsections eliciting demographic information, but parents of bilingual children additionally filled out information on language history and use.

Using the statistical package SPSS25, background information for the three groups was compared using one-way ANOVAs, which were followed up by Bonferroni tests for multiple comparisons. The results showed that the three groups were matched for chronological age ($F\ (2,54) = 2.66$, $p = .08$, $\eta^2 = 0.09$), although it should be noted that the age range was wide across all three groups. Furthermore, all the children in the sample came from mid-to-high SES families as determined by parental education. There were significant group differences with respect to the mother's years of education ($F\ (2, 53) = 5.92$, $p = .01$, $\eta^2 = .18$), yet no difference with respect to the father's years of education ($F\ (2, 53) = 0.48$, $p = .62$). The level of

2. These data partially overlap with those published at the BUCLD proceedings by Meir et al. (2022). For the current study, we collected a new dataset of monolingual Russian-speaking controls and increased the sample of monolingual Hebrew-speaking controls. This was done to ensure that all the children in the three groups were tested on production and using the same eye-tracking apparatus.

Table 2. Background information for the child participants (*M* (*sd*); range)

	BILING	MonoHE	MonoRU
Number of participants	19	19	19
Age in months	82 (18)	73 (16)	71 (13)
	53–106	49–103	57–96
Gender	12F/7M	8F/11M	12F/7M
Mother's education in years*	19 (3)	17 (2)	17 (1)
	15–25	12–22	16–22
Father's education in years*	16 (4)	17 (4)	16 (2)
	10–24	12–25	9–20

* An education level of 11–12 years corresponds to a high school diploma, 15–16 years – BA-level, 17–18 years – MA-level, 19 years and above doctorate and post-doctorate-level.

mother's education was significantly higher in bilinguals compared to both monolingual groups (BILING vs. MonoRu: $p = .017$; BILING vs. MonoHE: $p = .010$). The three groups did not differ with respect to gender distribution ($\chi 2(2) = 2.22$, $p = .32$), as determined by a Chi-Square test.

The bilinguals were exposed to their HL-Russian from birth, while the age of onset of SL-Hebrew varied ($M = 14$ months, $sd = 20$; range 0–48). The input available to the bilinguals from the mothers was mixed: 47.4% of the children were exposed exclusively to Russian; 47.4% had dual Russian-Hebrew exposure; 5.2% were exclusively exposed to Hebrew. The parental input also varied: 21.1% of the sample was exposed exclusively to Russian; 36.8% to dual Russian-Hebrew, 42.1% exclusively to Hebrew. We also asked the parents to rate their children's proficiency on a scale of 0–4 in the HL and the SL, with '0' corresponding to no knowledge of the language and '4' corresponding to excellent mastery of the language. As per parental ratings, proficiency in HL-Russian and SL-Hebrew was similar across the two languages, i.e., HL-Russian: $M = 3.32$ ($sd = 0.67$); SL-Hebrew: $M = 3.53$ ($sd = 0.61$). The paired-sample *t*-test confirmed that there was no significant difference between the scores in HL-Russian and SL-Hebrew ($t(19) = 0.94$, *p* = ~~0~~.36).

The study was approved by the IRB Committee of Bar-Ilan University. Informed parental consent and oral assent by each child were secured before the testing session(s) in accordance with the Declaration of Helsinki.

2.2 Design and materials

2.2.1 *Expressive vocabulary in Russian and Hebrew*

Expressive vocabulary scores were collected in order to obtain objective measures of children's proficiency, in addition to subjective parental proficiency ratings, described above. Expressive vocabulary was assessed via the production subtests of the *LITMUS Cross-Linguistic Task*, (*CLT*, Haman et al., 2015) targeting the production of nouns and verbs. Each subtask contains 31 items. We used the Russian (Ringblom & Dobrova, 2019) and the Hebrew (Altman et al., 2017) versions and evaluated expressive vocabulary in monolinguals and bilingual children. Table 3 presents the vocabulary scores for the three groups; bilinguals were tested in both languages.

Table 3. CLT expressive vocabulary scores *M* (*sd*); range

	BILING	MonoHE	MonoRU
N	19	19	19
Russian CLT:			
Nouns:	21 (5); 10–27	–	27 (2); 22–30
Verbs:	14 (5); 6–21	–	20 (4); 11–26
Total:	**35 (10); 16–47**	–	**47 (6); 36–54**
Hebrew CLT:			
Nouns	21 (7); 8–32	27 (2); 22–30	–
Verbs	15 (7); 2–25	21 (4); 15–26	–
Total	**36 (14); 10–54**	**48 (6); 38–56**	–

Note. CLT= Cross-Linguistic Task

A series of independent *t*-tests indicated that in both languages bilinguals scored significantly lower compared to their monolingual peers in HL-Russian (Nouns: $t(26.14) = 4.94$, $p < .001$; Verbs $t(26) = 4.06$, $p < .001$; total vocabulary: $t(29.37) = 4.81$, $p < .001$) and SL-Hebrew (Nouns: $t(22.11) = 3.34$, $p < .01$; Verbs: $t(26.87) = 2.94$, $p < .01$; total vocabulary: $t(24.21) = 3.19$, $p < .01$). The vocabulary production tasks confirmed that there were no significant differences between the two languages, as determined by paired-samples *t*-tests (all *p-values* > .05), confirming the parental ratings of the participants being balanced across their two languages.

2.2.2 *Production experiments: ACC case in Russian and Hebrew*

To test our RQ1, we conducted an elicitation production task in Russian (Janssen & Meir, 2019; Meir & Janssen, 2021) in which the child was asked to describe what s/he sees on the screen by saying *Ja vižu* ______ 'I see (target noun)'. The task elicits ACC case on 36 nouns. If the child failed to respond to the sentence with *Ja vižu* ______ 'I see ____', s/he was reminded to start the sentence with *ja vižu* 'I see'. This was done for each target noun to ensure that the ACC case was produced in the syntactic environment obligatory for ACC. In this study, we compared the production accuracy of nouns across two conditions: the marked conditions, i.e., the use of a dedicated ACC case inflection is required (e.g., *kukl-**a*** 'doll-NOM–*kukl-**u*** 'doll-ACC', *slon* 'elephant-NOM'–*slon-**a**'* 'elephant-ACC') and the unmarked condition, i.e., the ACC and NOM forms are homophonous (e.g., *mjod* 'honey-NOM = ACC')

The elicitation production task in Hebrew (Meir & Novogrodsky, 2021) evaluated the accuracy of production of the ACC marker *et* with definite nouns. The task elicits 20 noun phrases in indefinite (unmarked) and definite (marked) contexts in subject and object positions. The child was asked to describe visual stimuli following the experimenter's prompts (see (3)): in (3a), the child was expected to produce indefinite nouns in the object position, whereas in (3b), definite nouns in the object position with the ACC marker *et)*. It should be noted that in Hebrew, the ACC marker *et* and the definite marker *ha-* are often contracted to *"eta'* in oral speech; this response was marked as correct.

For Russian and Hebrew, production accuracy scores were calculated for marked and unmarked conditions. The production accuracy scores were analyzed using SPSS25.

(3)	**Experimenter's prompt**	**Target answer**	**Visual stimuli**
a.	Prompt:	*ma romi ciyera?* What did Romi draw?	
	Target:	*koxav.* star.INDEF	
b.	Prompt:	*ma yeš le- romi?* What does Romi have?	
	Target:	*kova ve sefer.* hat.INDEF and book.INDEF	

Prompt:	*ma romi asta?* What did Romi do?
Target:	*hi sama* ***et ha****-kova ba-kuvsa ve* ***et ha****-sefer ba-yalkut .* She put ACC DEF-hat into the box and ACC DEF-book into the backpack

2.2.3 *VWP experiments: On-line comprehension of the ACC case in Russian and Hebrew*

To test our RQ2, i.e., whether there are differences between Russian-Hebrew bilingual and Russian-speaking and Hebrew-speaking monolingual children in on-line comprehension of the two cues (ACC case and word order), we conducted two VWP eye-tracking experiments, in Russian and in Hebrew, in which Word Order (SVO vs. OVS) (4b–c) was manipulated. Each spoken sentence was preceded by a preamble (4a).

(4) a.

Russian:	*Èto*	*morkovka,*	*zajčik,*	*lisa.*
Hebrew:	*Ze*	*gezer,*	*arnav,*	*šual.*
	This	carrot	bunny	fox.

‘This is a carrot, bunny, fox.’

b.

Russian SVO:	*ser-yj*	*zajčik*	*sejčas*	*s’jest*	*morkovk-u.*
	grey-NOM	bunny-NOM	now	will-eat	carrot-ACC
Hebrew SVO:	*ha-arnav*	*ha-afor*	*yoxal*	*axšav*	*et ha-gezer*
	DEF-bunny	DEF-grey	will-eat	now	ACC DEF-*carrot*

‘The grey bunny will now eat the carrot.’

c.

Russian OVS:	*ser-ogo*	*zajčik-a*	*sejčas*	*s’jest*	*lis-a.*
	grey-ACC	bunny-ACC	now	will-eat	fox-NOM
Hebrew OVS:	*et ha-arnav*	*ha-afor*	*yoxal*	*axšav*	*ha-šual.*
	ACC DEF-bunny	DEF-grey	will-eat	now	DEF-fox

‘The grey bunny will now be eaten by the fox.’

The spoken sentences (4b–c) with the Word Order manipulation (the first independent variable) were crossed with two types of tasks represented in two different visual displays, namely, PIC-3 vs. PIC-2 (the second independent variable) (Figure 1).

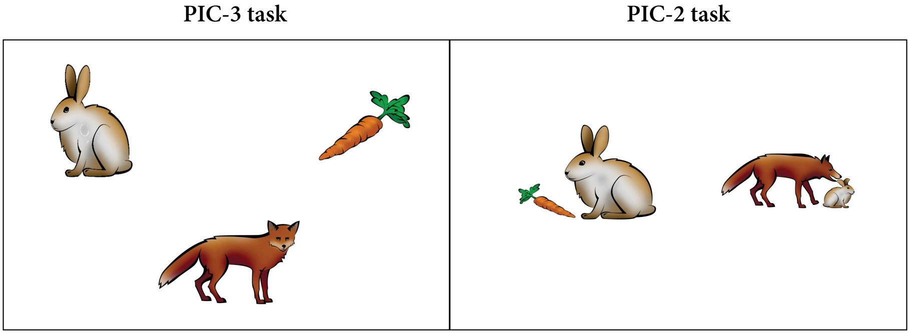


Figure 1. Task manipulation: PIC-3 vs. PIC-2

The PIC-3 task was introduced by Kamide et al. (2003) and further developed by Özge et al. (2019, 2022). It represents the classical 3-Referent context in which images of individual referents are not integrated into situations (e.g., isolated pictures of a bunny, fox, and carrot). Because our design intended to test the effects of word order found in these studies, we included the 3-Referent task into our experiment. However, the 3-Referent context (Figure 1, left panel) may not support the transitive event that is described in the auditory stimuli and could result in extra processing load for children that has nothing to do with cue weighing. The rationale for the PIC-2 task inclusion was two-fold. Firstly, we added the PIC-2 task because it involved the visual representation of the event in which the visual context requires forced-choice, thus helping to avoid the extra processing load, (Figure 1, right panel). In this task, the three individual referents are clearly engaged in a transitive *event*, with one of the referents (e.g., the bunny) being the *agent* of the eating event in one panel (paired with the sentence in 4b) and the *patient* of the eating event in the other (paired with the sentence in 4c). Secondly, the PIC-2 task was intended to replicate the results of the PIC-3 task previously implemented in testing case processing in German and Turkish (see Özge et al., 2019, 2022).

The PIC-3 task consisted of three steps: (1) Children saw the individual pictures of the three referents (Figure 1, left panel, PIC-3 task) while they listened to the preamble in (4a). (2) Then while the PIC-3 display remained on the screen, they heard the experimental sentence in either SVO word order (4b) or in OVS order (4c). Critically, it is during this step that their eye movements were recorded allowing us to establish at which point in the spoken sentences they would start looking at the bunny (agent or patient) and the fox (always agent). (3) Finally, immediately after that, they were presented with a single picture depicting the event (Figure 1, the left or the right from PIC-2 task) and performed a sentence-picture matching task. The PIC-2 task consisted of just two steps: (1) the same as in the PIC-3 task, and (2) the preamble display was replaced with the display with two events presented side-by-side (Figure 1, right panel, PIC-2 task) and the child performed a picture selection task while their eye movements were being recorded.

The experimental items were simple transitive sentences with two arguments (see 4b–c). In the SVO sentences, the first argument (NP1) was in the NOM case whereas the second (NP2) was in the ACC case. In the OVS sentences, NP1 was in the ACC case and NP2 was in the NOM case. In both languages, the verb occupied the medial position in the sentence, i.e., between NP1 and NP2. Beyond the obligatory constituents (NP1, NP2, and VERB), we also added an adjective or an adverb to modify NP1 in order to increase the duration of the predictive region. Note that in both languages, NP1 was modified by an adjective, but the order of the adjective placement varies between Russian and Hebrew. In Russian, the adjective appears before the NP and agrees in case marking with the NP, while in Hebrew, it appears after the NP, and it is not marked for case.

The sentences were divided into Regions of Interest (ROI): Adjective, first Noun (NP1), Adverb, Verb, and second Noun (NP2). The onset of each ROI was noted for each experimental sentence in each condition. Similar to the experiments in Turkish and German (Özge et al., 2019, 2022), the spoken sentences had plausible event structures which are likely to be encountered in real life (e.g., *bunnies eating carrots* and *foxes eating bunnies*). Once the NP2 is encountered in speech, the sentence could be processed based on the lexical items (bunny-carrot vs. bunny-fox). This procedure ensured children's successful comprehension performance. In the preamble (4a, Figure 1, left panel), the referents were introduced with their corresponding names to ensure that the children knew them.

The spoken sentences in the current study (4a–c) were adaptations of the German sentences from the previous experiments testing predictive ACC case processing in German (Özge et al., 2022). They were recorded by two female native speakers, one of Russian and the other of Hebrew, in a professional recording soundproof studio. All the sentences were recorded as a single audio file rather than cross-spliced to preserve natural intonation.

There were also five filler items in each task that were interspersed with the experimental items. The fillers included intransitive sentences (e.g., Russian: *Neposlushnaja loshad' klouna ubezhala* 'The clown's naughty horse ran away'; Hebrew: *ha-otobus ha-adom neecar al yad ha-bayit* 'The red bus stopped next to the house'). These sentences were not included in the current analysis. In total, the experiment included two practice items, 12 experimental items (six for each of the PIC-3 and PIC-2 tasks) and 10 fillers (five for each task). They were rotated through four lists in the Latin Square design.

2.3 Procedure

The children were tested individually at their kindergartens or community centers. The study was carried out under the supervision and in the presence of the first author. Monolingual children were tested in one session of 20–30 minutes, bilingual children were tested in two language sessions in the HL and the SL separated by at least one week. The testing order of the HL and the SL in bilingual children was counterbalanced. First, they completed vocabulary and ACC case production tasks, then participated in the eye-tracking experiment. Production tasks were audio-recorded for future off-line coding.

The participants' eye movements were recorded with an Eyelink Portable Duo eye-tracker (SR Research) at 1,000 Hz. Children sat in a comfortable chair in front of the eye tracker. The distance from the eyes to the camera was 60 cm, and 90 cm to the upper center of the monitor. The experiment started with the 9-point calibration. Each trial began with a drift correction dot (located in the screen center). After a successful drift correction, the picture(s) presentation began.

2.4 Eye-tracking data analysis

We extracted recorded eye-movement data using the *Eyelink Data Viewer* software to determine the average gaze position throughout each trial. We divided each trial into 20-ms bins in order to analyze the change in gaze over the course of the spoken sentence. Following Özge et al. (2019, 2022), we used the *agent preference* measure as a dependent variable in our analysis. Agent preference was calculated as the number of samples (for a given trial and ROI) in which the participant looked at the plausible agent, minus the number for the plausible patient (i.e., the looks to *the carrot* subtracted from the looks to *the fox* in the PIC-2 and in the PIC-3 tasks, in both word orders).

We used linear mixed-effects models to analyze the data via the *lme4* package (Bates et al., 2015) in R version 4.1.2 (R Core Team, 2021). To address our second and third research questions on similarities in on-line comprehension between monolingual and bilingual children and the effect of age, we fitted models with random and fixed effects evaluating agent preference in the predictive ROIs (Verb and Adverb). The random effects were random intercepts for participants and items. The fixed factors included Group (BILING vs. MonoRU)/(BILING vs. MonoHE), Condition (SVO vs. OVS), ROI (Verb Adverb), and Age. Contrast coding was used for the dichotomous predictor variables for ease of interpretation. We conducted the analysis separately for the PIC-3 and PIC-2 tasks because the isolated three-referent presentation (PIC-3) was visually very different from the two-event presentation (PIC-2) in the duration of the looks towards the Target. We included fixed effects and their interactions to determine whether there were differences across the two groups in each ROI. We used the *emmeans* and the *lsmeans* packages (Lenth et al., 2019) to conduct pairwise comparisons with Bonferroni corrections. Figures in the current study were generated using the *ggplot2* package (Wickham, 2009). Finally, for the last research question, we fitted another set of models only for the bilinguals evaluating the relationship between ACC production and predictive comprehension. The models also included random intercepts for participants and items. The fixed factors included Production (accuracy of ACC case production) and Condition (SVO vs. OVS).

3. Results

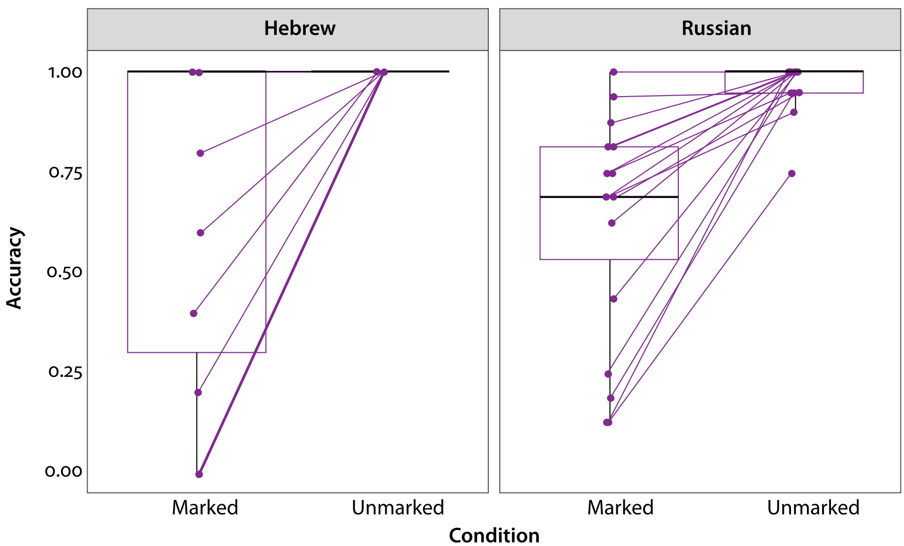
3.1 ACC case production accuracy in Russian and Hebrew

The ACC case production data were coded as the ratios of correct responses over the target nouns for marked and unmarked conditions, ranging between 0 and 1. The results revealed that the MonoRu and MonoHe controls showed a ceiling effect in the ACC case production for marked and unmarked forms (for both conditions $M = 1.00$, $sd = 0$). Thus, the production of the ACC case was error-free in monolingual controls.

The findings for the bilingual children were in sharp contrast to the monolinguals. Figure 2 demonstrates the scores across the two conditions in both languages in the BILING group. Bilinguals were nearly at ceiling on the unmarked forms in Russian ($M = 0.97$; $sd = 0.17$) and in Hebrew ($M = 1.00$; $sd = 0$). Yet in both languages they were less accurate with the marked forms: Russian ($M = 0.63$; $sd = 0.48$); Hebrew ($M = 0.68$, $sd = 0.47$). Using SPSS 25, we conducted a two-way ANOVA[3] with Language and Condition as repeated measures for the production data in bilinguals. This analysis aimed to compare the use of the ACC case in nouns that require ACC case marking versus those that do not. The results yielded a significant effect of Condition ($F\ (1,18) = 58.44$, $p < .001$, $\eta^2 = 0.77$), no effect of Language ($F\ (1,18) = .001$, $p = .97$) and no significant Condition*Language interaction ($F\ (1,18) = .10$, $p = .66$). Thus, the results show that overall, the BILING group was more accurate in the unmarked condition compared to the marked one.

3. For our production data, we conducted two-way ANOVAs to assess the influence of group and condition, as well as condition and language. This decision was made as the glmer models with random effects for item and participant failed to converge, and the glm model returned numerous warnings due to perfect separation in the data – specifically, ceiling performance in monolingual groups.

Subsequently, we applied separate two-way ANOVAs with Group as a between-subject variable and Condition as a within-subject variable to compare ACC case production in monolinguals and bilinguals in Russian and in Hebrew. In Russian, the analysis indicated a significant effect of Group (F (1,36) = 32.93, $p < .001$, $\eta^2 = 0.48$), a significant effect of Condition (F (1,36) = 33.88, $p < .001$, $\eta^2 = 0.49$) and a significant Group*Condition interaction (F (1,36) = 33.88, $p < .001$, $\eta^2 = 0.49$). Follow-up pairwise comparisons with Bonferroni adjustments indicated that the differences between monolinguals and bilinguals in the unmarked and marked conditions were significant ($p = .045$; $p < .001$, respectively). As for Hebrew, the analysis indicated a significant effect of Group (F (1,36) = 14.07, $p < .001$, $\eta^2 = 0.28$), a significant effect of Condition (F (1,36) = 14.07, $p < .001$, $\eta^2 = 0.28$) and a significant Group*Condition interaction (F (1,36) = 33.88, $p < .001$, $\eta^2 = 0.49$). Both groups showed error-free performance in the unmarked condition ($p = .76$), while in the marked condition, bilinguals were significantly less accurate ($p < .001$).



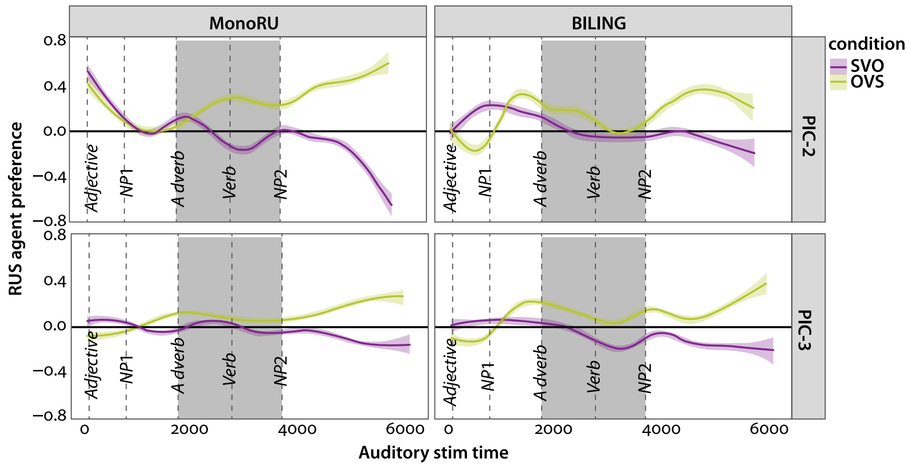
Note: Each dot represents one participant; the lines connect the scores for each participant across the two conditions. The upper line of the boxplots represents the first quartile (25%) and the lower line represents the third (75%), thus the box represents 50% of the data. The line inside the box represents the median.

Figure 2. ACC case-marking production in bilinguals in HL-Russian and SL-Hebrew

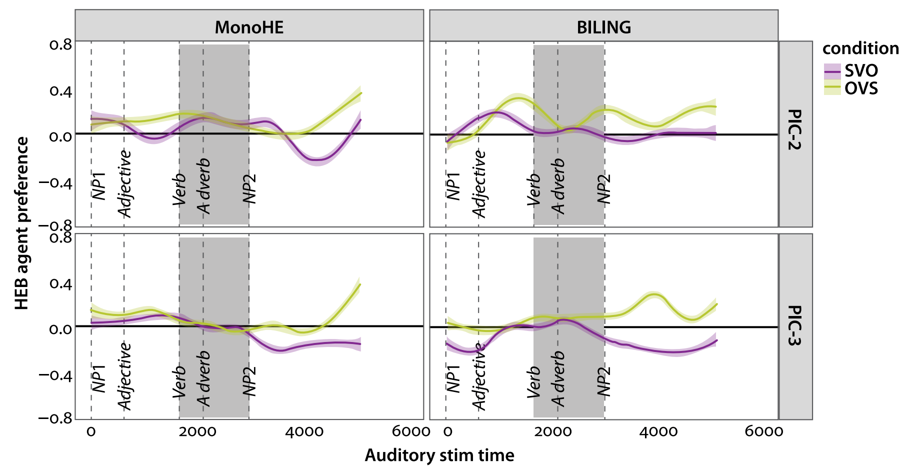
To sum up, while both groups of monolinguals showed an error-free performance on marked and unmarked ACC case conditions, bilinguals were less accurate compared to monolinguals and had specific difficulties with marking the ACC case in HL-Russian and in SL-Hebrew.

3.2 VWP experiments: ACC case comprehension in Russian and in Hebrew

The comparison of children's eye movements is presented in Figure 3 for the Russian and Hebrew data (BILING, MonoRU, MonoHE). The data show how the time course of agent preference changes through the course of the utterance. The looks above the zero line represent more looks to the subject (e.g., *fox*) and the looks below zero represent more looks to the patient (e.g., *carrot*). The grey area encompasses the two predictive ROIS (Adverb and Verb).



A. Russian



B. Hebrew

Figure 3. Agent preference (proportions of looks to the agent *fox* minus the patient *carrot*) in PIC-2 and PIC-3 tasks in bilingual and monolingual children. Top panels (A): Russian, Bottom panels (B): Hebrew

Table 4. Model summaries for PIC-2 and PIC-3 tasks in Russian and in Hebrew for the predictive ROIs (Verb and Adverb)

	Russian				Hebrew			
	PIC-2		PIC-3		PIC-2		PIC-3	
Predictors	*Estimates, SE, [95% CI]*	*p*	*Estimates, SE, [95% CI]*	*p*	*Estimates, SE, [95% CI]*	*p*	*Estimates, SE, [95% CI]*	*p*
(Intercept)	0.70, 0.26, [0.18, 1.21]	**.008**	−0.12, 0.14 [−0.39, 0.15]	.367	−1.13, 0.27 [−1.67, −0.60]	< **.001**	−0.36, 0.27 [−0.89, 0.18]	.193
Age	− .01, 0.00, [−0.01, −0.00]	**.003**	0.00, 0.00 [−0.00, 0.00]	.429	0.02, 0.00 [0.01, 0.02]	< **.001**	0.01, 0.00 [−0.00, 0.01]	.098
Group [MONO vs. BILING]	−0.38, 0.42, [−1.20, 0.44]	.365	0.10, 0.22 [−0.33, 0.53]	.648	1.74, 0.36 [1.03, 2.45]	< **.001**	0.28, 0.39 [−0.49, 1.05]	.482
Condition [SVO vs. OVS]	−0.72, 0.08, [−0.87, −0.58]	< **.001**	0.07, 0.05 [−0.02, 0.17]	.143	1.55, 0.12 [1.31, 1.79]	< **.001**	−0.54, 0.08 [−0.70, −0.39]	< **.001**
ROI [3 vs. 4]	−0.03, 0.02, [−0.07, 0.02]	.246	−0.02, 0.01 [−0.04, 0.01]	.290	−0.12, 0.03 [−0.18, −0.05]	**.001**	−0.14, 0.02 [−0.19, −0.10]	< **.001**
Age * Group [MONO vs. BILING]	0.01, 0.01, [−0.00, 0.02]	.261	−0.00, 0.00 [−0.01, 0.00]	.602	−0.02, 0.00 [−0.03, −0.02]	< **.001**	−0.01, 0.01 [−0.02, 0.00]	.302
Age * condition [SVO vs. OVS]	0.01, 0.00, [0.01, 0.01]	< **.001**	0.00, 0.00 [−0.00, 0.00]	.435	−0.02, 0.00 [−0.02, −0.02]	< **.001**	0.01, 0.00 [0.00, 0.01]	< **.001**
Group [MONO vs. BILING] * Condition [SVO vs. OVS]	0.51, 0.12, [0.27, 0.75]	< **.001**	−0.02, 0.08 [−0.17, 0.14]	.842	−1.27, 0.17 [−1.60, −0.94]	< **.001**	0.80, 0.12 [0.58, 1.03]	< **.001**
Group [MONO vs. BILING] * ROI [3vs 4]	−0.15, 0.03, [−0.21, −0.09]	< **.001**	−0.16, 0.02 [−0.20, −0.12]	< **.001**	0.08, 0.05 [−0.01, 0.17]	.083	0.14, 0.03 [0.08, 0.20]	< **.001**
Condition [SVO vs. OVS] * ROI [3 vs 4]	0.23, 0.03, [0.17, 0.29]	< **.001**	−0.06, 0.02 [−0.10, −0.02]	**.006**	−0.01, 0.05 [−0.10, 0.08]	.836	0.07, 0.03 [0.01, 0.13]	**.028**
Age * Group [MONO vs. BILING] * Condition [SVO vs. OVS]	−0.01, 0.00, [−0.01, −0.00]	< **.001**	0.00, 0.00 [−0.00, 0.00]	.213	0.02, 0.00 [0.01, 0.02]	< **.001**	−0.01, 0.00 [−0.01, −0.01]	< **.001**
Group [MONO vs. BILING] * Condition [SVO vs. OVS] * ROI [3 vs. 4]	−0.23, 0.05, [−0.32, −0.14]	< **.001**	0.12, 0.03 [0.06, 0.18]	< **.001**	0.08, 0.06 [−0.04, 0.21]	.190	−0.05, 0.04 [−0.14, 0.03]	.208

Note for the dichotomous variable Group – In Russian, MONO=MoRU; in Hebrew MONO=MoHE.
Note for the dichotomous variable ROI – In Russian: 3=Adverb; 4=Verb, in Hebrew 3=Verb, 4=Adverb (for more detail see 4b and 4c).

Table 4 summarizes coefficients of estimates (EST), standard errors (SE), 95% confidence interval for parameter values (CIs) and p-values for the linear mixed-effects models which were conducted separately for the PIC-2 and PIC-3 tasks for the Russian and Hebrew data. In Russian, the fitted model indicated significant three-way Group * Condition * ROI interactions for both tasks (PIC-2 and PIC-3) which were followed up by pairwise SVO-vs.-OVS contrast comparisons with Bonferroni corrections for multiple comparisons, for each group and each ROI separately. Following Brown (2021), when higher-order interactions are detected, main effects and lower-order interactions are ignored because they do not add to an understanding of the phenomenon. Pairwise SVO-vs.-OVS contrasts indicated that agent preference was significantly higher in OVS compared to SVO conditions in both groups and in both predictive ROI regions: for PIC-2 – MonoRU: Adverb: $EST=-0.04$, $SE=0.02$, $t=-1.67$, $p=.09$; Verb: $EST=-0.27$, $SE=0.02$, $t=-10.89$, $p<.0001$; BILING: Adverb: $EST=-0.13$, $SE=0.03$, $t=-4.85$, $p<.0001$; Verb: $EST=-0.13$, $SE=0.03$, $t=-4.84$, $p<.0001$; for PIC-3: MonoRU: Adverb: $EST=-0.11$, $SE=0.02$, $t=-6.74$, $p<.0001$; Verb: $EST=-0.05$, $SE=0.02$, $t=-3.05$, $p<.001$; BILING: Adverb: $EST=-0.20$, $SE=0.02$, $t=-10.93$, $p<.0001$; Verb: $EST=-0.25$, $SE=0.02$, $t=-14.70$, $p<.0001$.

The results were slightly different for Hebrew. A significant two-way Group * Condition interaction in the PIC-2 task was followed up by pairwise SVO-vs.-OVS contrast comparisons. The comparisons indicated that both groups showed agent preference in the PIC-2 task (MonoRU: $EST=-0.20$, $SE=0.03$, $t=-8.12$, $p<.0001$; BILING: $EST=-0.13$, $SE=0.02$, $t=-5.57$, $p<.0001$). In the PIC-3 task, there was a three-way Group * Condition * ROI interaction. Follow-up pairwise SVO-vs.-OVS contrasts indicated that agent preference was higher only in the BILING group, in both ROIs (Verb: $EST=-0.08$, $SE=0.02$, $t=-3.29$; $p=.001$; Adverb: $EST=-0.09$, $SE=0.02$, $t=-4.99$; $p<.0001$). In the monolingual Hebrew-speaking controls, agent preference was higher in the SVO condition ($EST=0.09$, $SE=0.03$, $t=3.78$, $p=.0002$) in the Verb region, yet no difference was observed in the Adverb region ($EST=0.03$, $SE=0.03$, $t=1.40$, $p=0.16$). Thus, predictive processing was observed in both groups in the PIC-2 task, yet in the PIC-3 task only in bilinguals.

3.3 The effect of age

Our model also tested to what extent chronological age is related to predictive processing in monolingual and bilingual children (RQ3). The models indicated the presence of Age * Group * Condition interactions (see Table 4) in Russian for PIC-2, in Hebrew for both PIC-2 and PIC-3. We followed up on the significant three-way Age * Group * Condition interactions focusing on the effect of Age in each group for each condition separately (see Figure 4) by applying linear mixed-effects models with Age as a predictor.

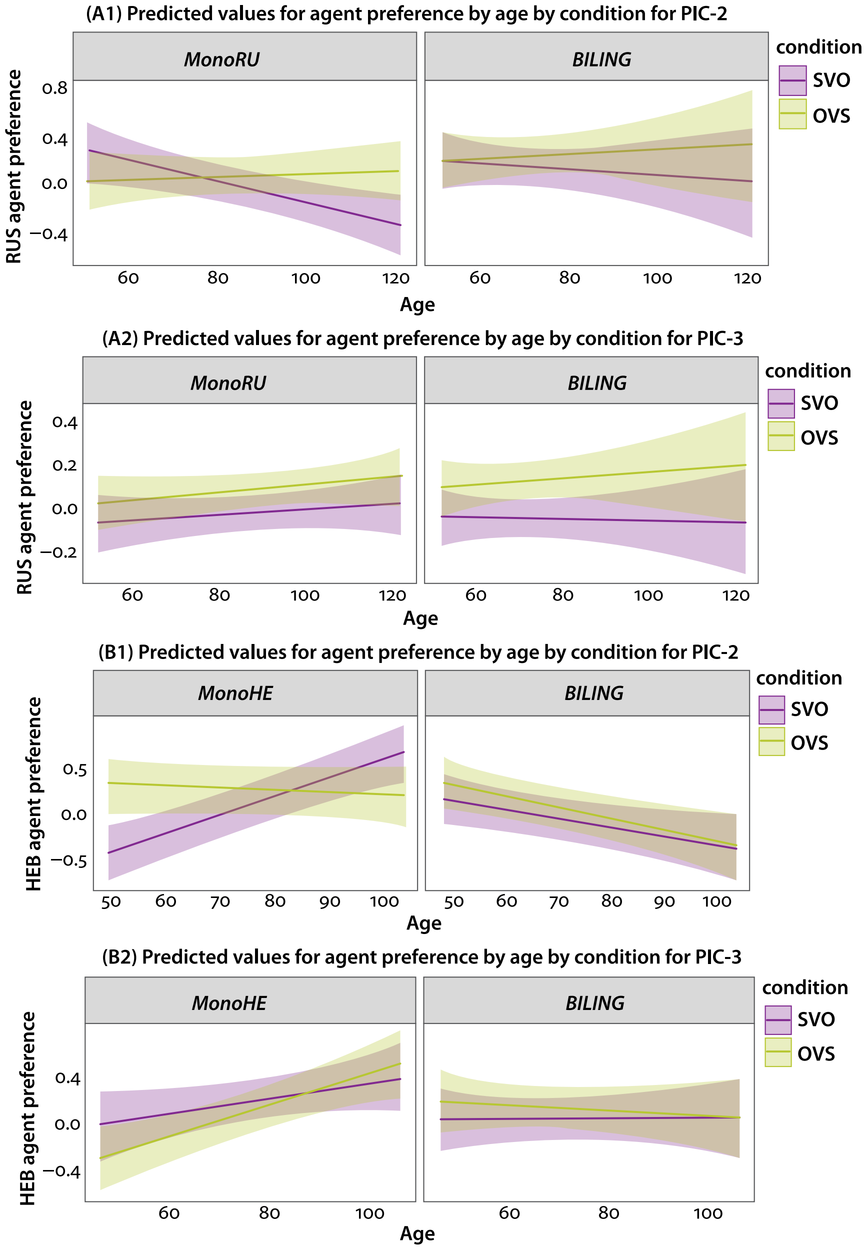


Figure 4. Proportions of agent preference as a function of age (in months) in the SVO and OVS sentences per group. Panel (A1): Russian (PIC-2), Panel (A2) Russian (PIC-3), Panel (B2): Hebrew (PIC-2), Panel (B2): Hebrew (PIC-3)

The follow-up analyses indicated that, in most cases, the effect of Age was not significant. However, in the Russian PIC-2 experiment, the effect of Age turned out to be significant only for the MonoRU group in the SVO condition (*EST* = 0.01, 95% *CI* [−0.02––0.00], *p* = .02), suggesting that, as expected, agent preferences decreased with age. Shifting to Hebrew, in the PIC-2 experiment, the effect of Age was significant in the SVO condition for MonoHE (*EST* = 0.01, 95% *CI* [−0.02––0.00], *p* = .008), while for bilinguals, it was significant in the OVS condition (*EST* = 0.02, 95% *CI* [0.0–0.03], *p* = .006). In monolinguals, agent preference increased in SVO, while in bilinguals, it decreased in OVS.

3.4 The role of the ACC case production in comprehension

In order to evaluate the link between production and comprehension (our RQ4), we additionally fitted models for Russian and Hebrew with agent preference score as the outcome variable, and the Condition * Production interaction as the fixed effect (see Table 5) with Participant as a random effect. We included ACC production scores as the index of Production. This was done only for bilinguals because the production data for monolinguals showed no variation due to the ceiling effect. We considered only the predictive regions in the two languages.

Table 5. Model summaries for bilingual children for Russian and in Hebrew for agent preference and Production (i.e., ACC case production) in the predictive ROIs (Verb and Adverb)

	Russian		**Hebrew**	
Predictors	***Estimates***	***p***	***Estimates***	***p***
(Intercept)	−0.01	0.887	−0.04	.615
Condition [SVO vs. OVS]	0.12	**< .001**	0.34	**< .001**
Production	−0.04	.750	0.06	.477
Condition [SVO vs. OVS] * Production	0.05	.252	−0.29	**< .001**

For Russian, the results indicated that the fixed predictor of Production and the Condition * Production interaction was not significant, suggesting that accuracy of production of the ACC case is not associated with predictive looks in Russian, as seen in Figure 5 (Panel A). In contrast, for Hebrew, the Condition * Production interaction turned out to be significant. Figure 5 (Panel B) demonstrates that with an increase in accuracy of ACC case-marking production, the agent preference decreases in the OVS condition in bilingual children. We elaborate on this surprising finding in the Discussion. A follow-up post-hoc test in Hebrew on the significant Condition * Production interaction indicated that the difference between OVS and SVO condition was significant ($EST = -0.11$, $SE = 0.01$, $z = -7.82$, $p < .0001$) when production score is controlled.

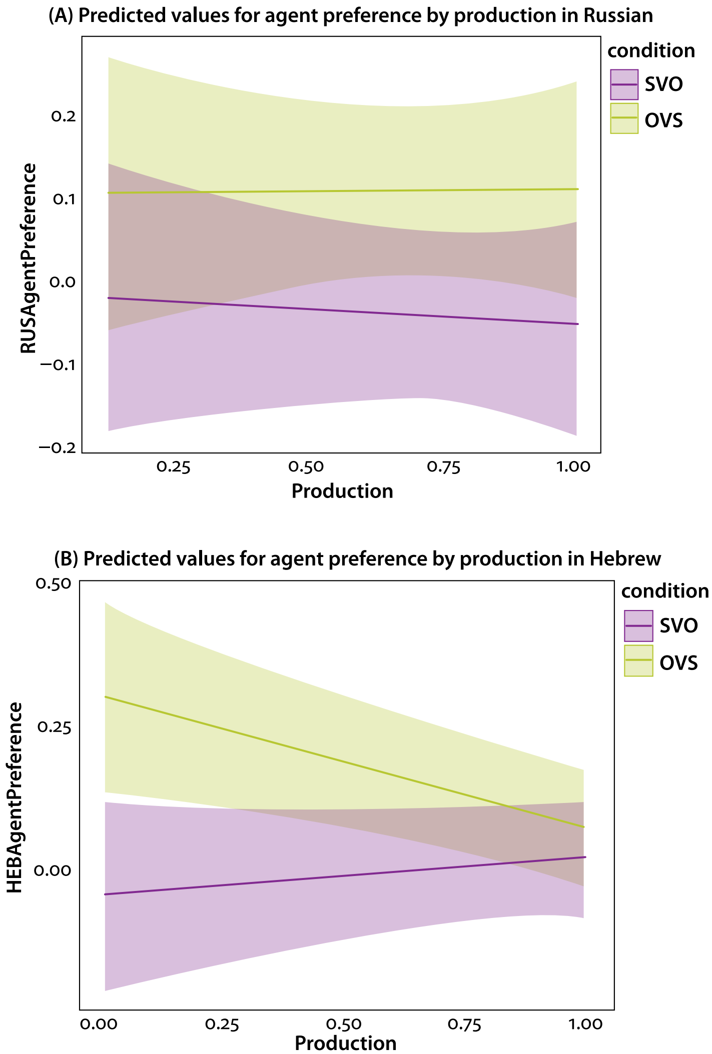


Figure 5. Proportions of agent preference in the SVO and OVS sentences as a function of ACC case production accuracy in bilingual children. Panel (A): Russian, Panel (B): Hebrew

4. Discussion

The current study investigated the production and the predictive use of the ACC case in 4-to-8-year-old Russian-Hebrew-speaking bilingual children who acquire Russian as their HL and Hebrew as their SL, and compared them to Russian- and Hebrew-speaking monolingual controls. This design allowed us to evaluate predictions of the *Unified Competition Model* (MacWhinney, 2012) in monolinguals and bilinguals. Both Russian and Hebrew use the ACC case-marking cue, yet the two languages vary with respect to its weight, as was found in previous studies. In Russian, case morphology is ubiquitous: every Russian noun and its modifying elements are marked morphologically for case which is critical for sentence interpretation. Case thus is the strongest cue for thematic role assignment for children and adults (Janssen & Meir, 2019; Kempe & MacWhinney, 1998; Sauermann & Gagarina, 2018). In Hebrew, the ACC case morphology is more limited: the ACC particle *et* is used only with definite nouns. For Hebrew-speaking monolingual adults, the ACC particle *et* is a reliable marker, but we know of no studies that tested predictive use of ACC case morphology in adult speakers. For children, the ACC particle *et* is less reliable, as reported in off-line studies (Frankel & Arbel, 1981; Sokolov, 1988).

Starting with our first research question that addressed production accuracy in monolingual and bilingual children, both monolingual groups, as expected, were at ceiling in both marked (i.e., Russian nouns requiring a dedicated case suffix *kukl-u* 'doll-ACC' vs. *kukl-a* 'doll-NOM' and in Hebrew *et ha-buba* 'ACC DEF-doll) and unmarked (i.e., Russian *stol* 'table-NOM = ACC', Hebrew indefinite nouns) ACC forms. In contrast, bilingual children were at ceiling only with the unmarked forms and produced the same number of non-target-like marked forms in both of their languages. Thus, our first research hypothesis was confirmed: on average, bilingual children were less accurate than monolinguals in their production of the ACC case marking in both languages.

Turning to our second research question that addressed on-line comprehension, a different picture emerged. Based on the *Unified Competition Model* (MacWhinney, 2012), we predicted the ACC case to be the strongest cue in Russian, but not in Hebrew. Our prediction was confirmed. In Russian-speaking monolinguals, there were more looks to the plausible agent in the OVS compared to the SVO sentences in the predictive regions (Adverb and Verb). This demonstrates that case marking in Russian is a reliable cue, and it is used immediately in predicting the upcoming information. In contrast, Hebrew-speaking monolinguals, despite error-free production, did not rely on the ACC case marking *et* in the PIC-3 task, yet did so in the PIC-2 task. This suggests that monolingual Hebrew-speaking children's reliance on case morphology to predict the upcoming thematic roles is weaker than in Russian-speaking peers. The Russian results confirmed that the predictive use of a reliable cue does not depend on the task. But predictive processing can be weakened for some children if the task does depict separate referents, as demonstrated for the Hebrew findings. This methodological difference between the two tasks deserves further exploration in future studies.

How can we explain the difference between Russian- and Hebrew-speaking monolingual children's results? We suggest that it stems from inherent properties of the cue weights in the two languages. In Russian, the ubiquitous nature of the ACC case and its doubling (i.e., on the adjective and on the noun) make it a very strong cue in contrast to a single cue and its limited use in Hebrew (only in definite contexts). To level the playing field between the two languages, we could remove the modifiers and only use sentences in Russian with unmodified NPs (e.g., *Zajčik-a sejčas s'jest lisa* 'Rabbit-ACC now will eat fox-NOM). Another option afforded by the Russian syntax is sentences with split construction (e.g., *Ser-ogo sejčas s'jest zajčik-a lisa* 'Grey-ACC now will eat rabbit-ACC fox-NOM) (Sekerina & Trueswell, 2011). These two options would allow us to disentangle the number of ACC case markings from their individual strength. We leave them for future investigation.

As far as predictive abilities in bilingual children are concerned, we outlined several scenarios based on the *Unified Competition Model* (MacWhinney, 2012). The present study brought evidence for the null hypothesis and partially for the transfer hypothesis. Bilinguals were indistinguishable from their monolingual peers in on-line comprehension in their HL-Russian. In SL-Hebrew, they were similar to their monolingual peers, but only in the PIC-2 task that highlights the event interpretation. In the PIC-3 task, bilingual children even showed an advantage over their monolingual Hebrew-speaking peers. A potential explanation for this result lies in the partial transfer of the strong case-marking cue from HL-Russian to SL-Hebrew that is language-pair specific: it is a result of cue weight in Russian affecting Hebrew. Weaker cues in one language, i.e., Hebrew, are reinforced by stronger cues in the other language, i.e., Russian.

Our third research question evaluated the role of age in predictive processing. Our study included monolingual and bilingual children aged 4–8. Previous research pointed to differential effects of age in monolingual and bilingual children (Paradis, 2023). In monolingual children, age stands out as a robust predictor of language abilities. However, in bilingual children, diverse outcomes may emerge: with increasing age, they might experience fossilization and/or attrition in their HL. In the current study, the effects of age showed mixed results. In most conditions and groups, age effects were not significant, which differs from the results reported in Aumeistere et al. (2022). However, in both monolingual groups (MonoRU and MonoHE), significant age effects were observed for SVO sentences but not for OVS sentences. In monolingual children, in Russian, as expected, agent preferences decreased in SVO sentences, while in Hebrew, surprisingly, they increased. In bilinguals, in Hebrew, as expected, age preferences decreased in OVS sentences. Consequently, it appears that input factors could potentially have a more substantial impact than age. Future studies should thoroughly investigate the influence of input and age in bilingual predictive processing.

Finally, our fourth research question addressed a connection between production and predictive processing in on-line comprehension. The *Prediction-by-Production Account* (Pickering & Gambi, 2018; Pickering & Garrod, 2013) *posits* tight links between production and prediction, whereas the *Missing Surface Inflection Hypothesis* (Prévost & White, 2000) suggests that despite lower production skills, bilinguals might be on par with monolinguals in comprehension. We propose that our two pieces of evidence combined point to a likely production-prediction dissociation and support the *Missing Surface Inflection Hypothesis*. First, the bilinguals were on par with the Russian monolinguals in predictive processing and showed an advantage compared to their monolingual peers in Hebrew, despite being significantly lower in accuracy of production of the ACC case. These results are in line with previous studies showing that despite a considerable production error rate, bilinguals are sensitive to ungrammaticalities, i.e., they show monolingual-like patterns of integration of morphosyntactic information (Blom et al., 2016; Chondrogianni & Marinis, 2012). Second, our statistical analysis revealed no associations between production and on-line comprehension for HL-Russian, and in SL-Hebrew predictive processing was unexpectedly negatively associated with production. Negative associations between production accuracy in Hebrew and predictive processing in bilinguals might indicate a U-shaped curve in language acquisition. Between the ages of 4 and 6, bilingual children's production of the ACC case in Hebrew is still fragile whereas the comprehension and sensitivity to ungrammaticalities is intact, consistent with the *Missing Surface Inflection Hypothesis*. Future studies should investigate production and prediction in monolingual Hebrew-speaking children of different ages in more detail.

5. Future research and conclusions

Although our study provided important insights into the predictive use of morphology in monolingual and bilingual children, it is not without limitations. First, our sample size was small, thus, the conclusions must be treated with caution. In addition, although the effect of age in predictive processing was addressed in the current study, future studies need to investigate it in more detail by comparing children of different age groups, i.e., 4–5-, 7–8-, and 10–11-year-olds. Furthermore, in future studies the interaction of the cues in bilinguals should be evaluated relative to children's language dominance in order to further assess the role of production in predictive processing. The children in our sample were balanced bilinguals, yet more information is needed on unbalanced bilinguals. Future studies should determine whether in unbalanced bilinguals, i.e., those children who are dominant in SL-Hebrew, the transfer of processing strategies will be from the dominant SL-Hebrew onto their weaker HL-Russian, and as a result the observed advantage for bilinguals might disappear.

To conclude, on the theoretical side, our study was the first to confirm the predictions of the *Unified Competition Model* (Macwhinney, 2012) and the *Missing Surface Inflection Hypothesis* (Prévost & White, 2000) using the Visual World eye-tracking Paradigm. The results showed cross-linguistic differences in the strength of the case-marking cue and its predictive use in sentence processing. For bilingual child processing, the findings indicated that bilingual children were on par with their monolingual peers in the HL in the use of ACC case morphology, and they showed similar or enhanced performance in the SL (as found in the PIC-3 task for Hebrew) despite their lower production accuracy.

On the methodological side, the study points to the role of the task (isolated referents versus helpful event depiction) that can subtly affect the performance of the children in the PIC-2 and PIC-3 tasks, resulting in better performance on the PIC-2 task. The differences in the task should be controlled in future studies.

Funding

We gratefully acknowledge the financial support provided by the United States-Israel Binational Science Foundation (BSF), Grant Number 2022245, for the project led by Principal Investigators [Natalia Meir & Irina A. Sekerina], titled "Language processing in bilingual children in L1 and L2: Prediction and integration of morphosyntactic information.

Data availability

The materials, data, and analysis script for this study can be retrieved from: https://osf.io/fapdm/?view_only=74c98955deef4b15824d13d3af4a7bd3

Competing interests

The authors declare no competing interests.

Acknowledgements

The authors would like to thank all the participants of the study for their cooperation. The authors extend gratitude to the anonymous reviewers and editors for their invaluable and constructive feedback, which has greatly contributed to enhancing the quality of the manuscript.

References

doi Abutbul-Oz, H., & Armon-Lotem, S. (2022). Parent Questionnaires in Screening for Developmental Language Disorder Among Bilingual Children in Speech and Language Clinics. *Frontiers in Education*, *7*.

Allen, S. (2014). Argument Structure. In P.J. Brooks & V. Kempe (Eds.), *Encyclopedia of Language Development*. (pp. 18–19). Sage Publications.

doi Altman, C., Goldstein, T., & Armon-Lotem, S. (2017). Quantitative and qualitative differences in the lexical knowledge of monolingual and bilingual children on the LITMUS-CLT task. *Clinical Linguistics & Phonetics*, *31*(11–12), 931–954.

doi Aumeistere, A., Bultena, S., & Brouwer, S. (2022). Wisdom comes with age? The role of grammatical gender in predictive processing in Russian children and adults. *Applied Psycholinguistics*, *43*(4), 867–887.

Bailyn, J.F. (2012). *The Syntax of Russian*. Cambridge University Press.

doi Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

doi Berman, R.A. (1981). Regularity vs anomaly: the acquisition of Hebrew inflectional morphology. *Journal of Child Language*, *8*(2), 265–282.

doi Biran, M., & Ruigendijk, E. (2015). Do case and gender information assist sentence comprehension and repetition for German-and Hebrew-speaking children? *Lingua*, *164*, 215–238.

doi Blom, E., Chondrogianni, V., Marinis, T., & Vasić, N. (2016). The acquisition of verbal paradigms in Dutch and Greek L2 children: Cross-linguistic differences and inflectional defaults. *International Journal of Bilingualism*, *20*(4), 386–402.

Bosch, J.E. & Foppolo, F. (2022). Predictive processing of grammatical gender in bilingual children: The effect of cross-linguistic incongruency and language dominance. *Lingue e Linguaggio*, *21*(1), 5–27.

doi Bosch, J.E., & Foppolo, F. (2023). Prediction during spoken language processing in monolingual and multilingual children: Investigating the role of literacy. *Linguistic Approaches to Bilingualism*.

doi Bosch, J.E., Chailleux, M.A., Yee, J., Guasti, M.T., & Arosio, F. (2022). Prediction on the basis of gender and number in Mandarin-Italian bilingual children. In Ayoun, D. (Ed.), *The acquisition of gender* (pp. 243–272). John Benjamins Publishing.

doi Brouwer, S., Sprenger, S., & Unsworth, S. (2017). Processing grammatical gender in Dutch: Evidence from eye movements. *Journal of Experimental Child Psychology*, *159*, 50–65.

doi Brown, V.A. (2021). An introduction to linear mixed-effects modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1).

Cejtlin, S. N. (2009). *Ocherki po Slovoobrazovaniyu i Formoobrazovaniyu v Detskoi Rechi.* Znak.

Chrabaszcz, A., Ladinskaya, N., & Lopukhina, A. (2023). Acquisition of Russian noun case by bilingual children: lexical cues to case assignment in real and novel words. *Language Acquisition*, 1–29.

Chondrogianni, V., & Marinis, T. (2012). Production and processing asymmetries in the acquisition of tense morphology by sequential bilingual children. *Bilingualism: Language and Cognition*, *15*(1), 5–21.

Chondrogianni, V., & Schwartz, R. G. (2020). Case marking and word order in Greek heritage children. *Journal of Child Language*, *47*(4), 766–795.

Danon, G. (2001). Syntactic definiteness in the grammar of Modern Hebrew. *Linguistics*, *39*(6), 1071–1116.

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German Childrens Comprehension of Word Order and Case Marking in Causative Sentences. *Child Development*, *79*(4), 1152–1167.

Ferreira, F., & Chantavarin, S. (2018). Integration and prediction in language processing: A synthesis of old and new. *Current Directions in Psychological Science*, *27*(6), 443–448.

Frankel, D. G., Amir, M., Frenkel, E., & Arbel, T. (1980). A developmental study of the role of word order in comprehending Hebrew. *Journal of Experimental Child Psychology*, *29*(1), 23–35.

Frankel, D. G., & Arbel, T. (1981). Developmental changes in assigning agent relations in Hebrew: The interaction between word order and structural cues. *Journal of Experimental Child Psychology*, *32*(1), 102–114.

Friedmann, N., & Shapiro, L. P. (2003). Agrammatic comprehension of simple active sentences with moved constituents: Hebrew OSV and OVS structures. *Journal of Speech Language and Hearing Research*, *46*, 288–297.

Gagarina, N., & Klassert, A. (2018). Input Dominance and Development of Home Language in Russian-German Bilinguals. *Frontiers in Communication*, *3*.

Grimm, A., Müller, A., Hamann, C., & Ruigendijk, E. (2011). *Production-comprehension asymmetries in child language.* Walter de Gruyter.

Hacohen, A., Kagan, O., & Plaut, D. (2021). Differential object marking in modern Hebrew: Definiteness and partitivity. *Glossa: A Journal of General Linguistics*, *6*(1).

Haman, E., Łuniewska, M., & Pomiechowska, B. (2015). Designing Cross-Linguistic Lexical Tasks (CLTs) for Bilingual Preschool Children. In Armon-Lotem, S., de Jong, J., & Meir, N. (Eds.), *Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment* (pp. 196–240). Multilingual Matters.

Hendriks, P., & Koster, C. (2010). Production/comprehension asymmetries in language acquisition. *Lingua*, *120*(8), 1887–1897.

Ionin, T., Goldshtein, M., Luchkina, T., & Styrina, S. (2023). Who did what to whom, and what did we already know? Word order and information structure in heritage and L2 Russian. *Linguistic Approaches to Bilingualism*, *13*(3), 343–371.

Janssen, B., & Meir, N. (2019). Production, comprehension and repetition of accusative case by monolingual Russian and bilingual Russian-Dutch and Russian-Hebrew-speaking children. *Linguistic Approaches to Bilingualism*, *9*(4–5), 736–765.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*(1), 133–156.

Karaca, F., Brouwer, S., Unsworth, S. & Huettig, F. (2021). Prediction in bilingual children: The missing piece of the puzzle. In Kaan, E. & Grüter, T. (Eds.), *Prediction in Second Language Processing and Learning* (pp. 116–137). John Benjamins.

Kempe, V., & MacWhinney, B. (1998). The acquisition of case marking by adult learners of Russian and German. *Studies in Second Language Acquisition*, *20*(4), 543–587.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59.

Lemmerth, N., & Hopp, H. (2019). Gender processing in simultaneous and successive bilingual children: Cross-linguistic lexical and syntactic influences. *Language Acquisition*, *26*(1), 21–45.

Lenth, R., Singmann, H., Love, J., Buerkner, P. & Herve, M. (2019). *Package 'emmeans'*. https://cran.r-project.org/web/packages/emmeans/index.html

MacWhinney, B. (2012). The logic of the unified model. In Mackey, A. & Gass, S. M. (Eds.), *The Routledge Handbook of Second Language Acquisition*. (pp. 211–227). Routledge. http://psyling.psy.cmu.edu/papers/years/2012/logic.pdf

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of verbal learning and verbal behavior*, *23*(2), 127–150.

MacWhinney, B., & Bates, E. (1989). *The Crosslinguistic Study of Sentence Processing*. Cambridge University Press.

Martin, C. D., Branzi, F. M., & Bar, M. (2018). Prediction is production: The missing link between language production and comprehension. *Scientific Reports*, *8*(1), 1079.

Meir, N., Avramenko, M., & Verkhovtceva, T. (2021). Israeli Russian: Case morphology in a bilingual context. *Russian Journal of Linguistics*, *25*(4), 886–907.

Meir, N., Parshina, O., & Sekerina, I. A. (2020). The Interaction of Morphological Cues in Bilingual Sentence Processing: An Eye-Tracking Study. In Megan M. Brown & Alexandra Kohut (Eds.), *BUCLD 44: Proceedings of the 44th annual Boston University Conference on Language Development*. Somerville (pp. 376–389). Cascadilla Press.

Meir, N., & Janssen, B. (2021). Child Heritage Language development: An interplay between cross-linguistic influence and language-external factors. *Frontiers in Psychology*, *12*.

Meir, N., & Novogrodsky, R. (2021). Referential expressions in monolingual and bilingual children with and without Autism Spectrum Disorder (*ASD*): A study of informativeness and definiteness. *Journal of Child Language*, *50*(2), 215–244.

Özge, D., Kornfilt, J., Maquate, K., Küntay, A. C., & Snedeker, J. (2022). German-speaking children use sentence-initial case marking for predictive language processing at age four. *Cognition*, *221*, 104988.

Özge, D., Küntay, A., & Snedeker, J. (2019). Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition*, *183*, 152–180.

Paradis, J. (2023). Sources of individual differences in the dual language development of heritage bilinguals. *Journal of Child Language*, *50*(40), 793–817.

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*(10), 1002–1044.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105–110.

Pickering, M. J., & Garrod, S. (2013). How tightly are production and comprehension interwoven? *Frontiers in Psychology*, *4*. Art. 238.

Plaut, D., & Hacohen, A. (2022). The acquisition of Hebrew differential object marking: Between production and comprehension. *Journal of Child Language Acquisition and Development-JCLAD*, *10*(4), 627–652.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, *16*(2), 103–133.

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.r-project.org/index.html

Ringblom, N., & Dobrova, G. (2019). Holistic constructions in heritage Russian and Russian as a second language: Divergence or delay? *Scando-Slavica*, *65*(1), 94–106.

Sauermann, A., & Gagarina, N. (2018). The roles of givenness and type of referring expression in the comprehension of word order in Russian-speaking children. *Linguistics Vanguard*, *4*(s1).

Schwartz, M., & Minkov, M. (2014). Russian Case System Acquisition among Russian-Hebrew Speaking Children. *Journal of Slavic Linguistics*, *22*(1), 51–92. https://www.jstor.org/stable/24602167.

Sekerina, I. A., & Trueswell, J. C. (2011). Interactive processing of contrastive expressions by Russian children. *First Language*, *32*(1–2), 63–87.

Sekerina, I. A. (2014). Visual World Eye–Tracking Paradigm. In P. J. Brooks and V. Kempe, (Eds.) *Encyclopedia of Language Development*. (pp. 658–658). SAGE Publications.

Shvedova, N. Y. (1980). *Russkaya grammatika [Russian grammar]*. Institute of the Russian Language, Russian Academy of Sciences. (in Rus.)

Slioussar, N., & Makarchuk, I. (2022). SOV in Russian: A corpus study. *Journal of Slavic Linguistics*, *30*, 1–14. http://ojs.ung.si/index.php/JSL/article/view/98

Sokolov, J. L. (1988). Cue validity in Hebrew sentence comprehension. *Journal of Child Language*, *15*(1), 129–155.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer.

Address for correspondence

Natalia Meir
Department of English Literature and Linguistics & The Gonda Multidisciplinary Brain Research Center
Bar-Ilan University
Max ve- Anna Webb
5290002 Ramat Gan
Israel

natalia.meir@biu.ac.il
https://orcid.org/0000-0001-9426-811X

Co-author information

Olga Parshina
Department of Psychology
Middlebury College
oparshina@middlebury.edu
https://orcid.org/0000-0002-6476-4961

Irina A. Sekerina
Department of Psychology and Ph.D. Program in Linguistics
College of Staten Island and The Graduate Center
The City University of New York
irina.sekerina@csi.cuny.edu
https://orcid.org/0000-0003-3859-3000

Publication history

Date received: 16 December 2022
Date accepted: 5 January 2024