Research in Generative Grammar
RGG 2025.15
Rivista di Grammatica Generativa

How do L2 learners of LIS process code-blending?

Caterina Donati[1], Carlo Geraci[2], Beatrice Giustolisi[3], Giulia Multineddu[4] & Chiara Branchini[4]

[1]University Paris Cité, [2]Institut Jean-Nicod (ENS, EHESS, CNRS), PSL Research University, [3]University of Milano-Bicocca, [4]Ca' Foscari University of Venice

Abstract. Semantic categorization tasks indicate a processing advantage of lexical bimodal-bilingual stimuli (simultaneous production of a sign and a spoken language) over unimodal stimuli (only signed or only spoken language) in the population of adult native signers (Children of Deaf Adult, CODAs). This finding has been reported for American Sign Language-English (Emmorey et al. 2012), Italian Sign Language (LIS)-Italian, and French Sign Language-French (Giustolisi et al. 2024). Such an advantage is not found when stimuli are sentences (Giustolisi et al. 2024). In this work, we report two experiments in LIS-Italian to explore whether these findings reported for the CODA population are also found in second language learners of LIS. Experimental results show that L2 learners do not have the same advantage as CODAs at the lexical level, but that frequent L2 users of LIS have a similar behavior in the processing of bimodal-bilingual sentences as CODAs.

Cite as: Donati, Caterina, Carlo Geraci & Beatrice Giustolisi & Giulia Multineddu & Chiara Branchini. 2025. How do L2 learners of LIS process code-blending?. *RGG*, 47(15). pp. 1–19. Lingbuzz Press.

1 Introduction. Code blending and its advantage in comprehension

Bilingual individuals often use both their languages within the same utterance (code-mixing). A natural and peculiar behavior of bimodal bilinguals, i.e., individuals who are proficient in a spoken language and in a sign language, is that, when engaged in code-mixing, they tend not to successively alternate their two languages, as unimodal bilinguals do, but rather prefer to use the two languages simultaneously (Emmorey, Petrich & Gollan 2012: 199). This simultaneous mixing, called code-blending, in opposition to a sequential one (code-switching), has been observed across all populations of bimodal bilinguals, both Deaf and hearing and represents a hallmark of this specific type of bilingualism (American Sign Language (ASL)/English, Bishop & Hicks 2005; Dutch Sign Language (NGT)/Dutch, Baker & Van Den Bogaerde 2008; Quebec Sign Language (LSQ)/French, Petitto et al. 2001; Brazilian Sign Language (LIBRAS)/Portuguese, De Quadros 2018; Italian Sign Language (LIS)/Italian, Donati & Branchini 2013; Branchini & Donati 2016; Finnish Sign Language (FinSL)/Finnish, Kanto, Huttunen & Laakso 2013; etc.).

On the comprehension side, it has been shown that lexical code-blendings of the type in (1c) tend to be processed faster than their unimodal equivalents in (1a) and (1b) by CODA bilinguals (hearing individuals born into deaf signing families who acquired a sign and a spoken language from birth).

(1) a. ananas (Italian)
b. PINEAPPLE (LIS)
c. ananas (code-blend Italian part)
PINEAPPLE (code-blend LIS part)

(Giustolisi et al. 2024 – online materials)

In a number of experiments, CODAs were shown to be faster in performing a simple semantic categorization task (e.g., Is the item edible?) when a lexical input was given bimodally (code-blending condition) than when it was given unimodally (either speech only or sign only conditions). The first results in this direction were reported in Emmorey, Petrich & Gollan (2012) with ASL/English bilinguals and were replicated in Giustolisi et al. (2024) with LIS/Italian bilinguals and LSF/French bilinguals.

Crucially, the advantage of code-blending did not emerge only with respect to the sign language (effect found with American and Italian CODAs), which is known to be the non-dominant language in hearing bilinguals, but also with respect to the spoken language, which is their dominant language (effect found in all three populations, American English, French, and Italian-speaking CODAs). These results suggest that the simultaneous elaboration of two congruent lexical items stemming from two languages and modalities is costless for early bimodal hearing bilinguals. A possible explanation for this effect is to attribute it to a general cognitive phenomenon known as the redundant signals effect, namely the fact that subjects respond more quickly when two stimuli with congruent information are presented in different modalities compared to when just one stimulus is presented (Miller 1986 and Emmorey, Petrich & Gollan 2012).

A similar experiment, but involving stimuli consisting of full sentences, as in (2, 3), and a truth value judgement task (e.g., Is the sentence generally true?) did not show the same advantage for code-blending when the bimodal condition was compared to the spoken language condition, neither in LIS/Italian, where the two simultaneous sentences had different word orders, as in (2), nor in French/LSF, where the two sentences were systematically aligned (3).

(2) a. I leoni mangiano le brioche (Italian)
the lions eat the croissants
b. LIONS CROISSANT EAT (LIS)
c. I leoni mangiano le brioche (code-blend Italian part)
LIONS CROISSANT EAT (code-blend LIS part)
'Lions eat croissants.'

(Giustolisi et al. 2024: 5)

(3) a. Les lions mangent les croissants (French)
the lions eat the croissants
b. LIONS EAT CROISSANT (LSF)
Les lions mangent les croissants (code-blend French part)
LIONS EAT CROISSANT (code-blend LSF part)
'Lions eat croissants'

(Giustolisi et al. 2024: 4)

Specifically, French CODAs showed no difference in the code-blended/French comparison (3b) vs. (3a); while Italian CODAs were quicker in processing the Italian only stimuli compared with the code-blended stimuli (code-blended/Italian comparison, (2c) vs. (2a)). This second finding suggested a processing cost at the sentential level, probably generated by the incongruent word order of the two languages in the blended stimuli (see Giustolisi et al. 2024 for a more extensive discussion). On the other hand, when comparing code-blended and sign language only stimuli, Giustolisi et al. (2024) found a significant advantage for the code-blended condition both for the code-blended/LSF pair (3b) vs. (3b) and for the code-blended/LIS pair (2c) vs. (2b). Language dominance could be the reason for this advantage. In fact, adult CODAs have been shown to be dominant in their spoken language compared to their sign language (e.g., Emmorey, Petrich & Gollan 2013).

One question that remained open from these previous studies is whether these effects are also found in late L2 learners of a sign language. Emmorey, Petrich & Gollan (2012)'s original study did include late bimodal ASL/English bilinguals (i.e., hearing people who have been exposed to ASL after the sensitive period) and found preliminary evidence of a possible advantage for bimodal stimuli over English only stimuli in the lexical task also for late learners, but only with high frequent words. The late ASL/English bilinguals (N = 25) that were included in that study used ASL on a daily basis and 16 of them were professional interpreters. The study did not include more complex sentential stimuli.

The aim of this paper is to deepen our understanding of the code-blended lexical facilitation and the absence of facilitation in the sentential task observed in Italian CODAs by investigating whether hearing late L2 learners of LIS show effects similar to those observed in CODAs. We will also investigate whether these effects are dependent on the frequency of use of LIS. The answer to these questions is very relevant as it might disclose whether this pattern is specific of native bilinguals (CODAs) or whether they can potentially emerge in all bimodal users.

For this purpose, we administered the same two (lexical and sentential) tasks as in Giustolisi et al. (2024) to L2 learners of LIS controlling for the frequency of use of LIS and the numbers of years they have been exposed to LIS. On the one hand, since this population is likely to be biased by the spoken language even more than CODAs, we expect an advantage of code-blending over the LIS only condition. On the other hand, if a certain amount of experience in a language in the visual modality is enough to trigger lexical facilitation, we also expect to find an advantage of the blended condition over the Italian only condition in the lexical experiment. As for the comparison of code-blended stimuli with Italian only stimuli in the sentential task, we expect to find either a disadvantage for the bimodal condition, similar to what was found in Giustolisi et al. (2024), or no difference between the two conditions, if the sign language dimension of the stimuli is ignored.

2 Methods

In the present work, we used the same semantic decision task (LEX-I) and the same truth value judgement task (SEN-I) employed with Italian CODAs participants in Giustolisi et al. (2024). To answer our new research questions, the population in the present work consists of hearing signers exposed to LIS as L2 in adulthood.

2.1 Participants

Participants were invited to take part in our research through personal contacts. The participants were or had been LIS students at Ca' Foscari University of Venice and some of them were professional LIS interpreters. All participants had at least 4 years of LIS use. A subpart of participants (33) were still students enrolled in a master's degree program in either "Language Sciences" or "Translation and Interpretation" at Ca' Foscari University of Venice and were about to obtain their degree. The Master students had studied LIS during their Bachelor and Master, for a total of four years for the MA in "Language sciences" and a total of five years for the MA in "Translation and Interpretation". The level of knowledge of LIS achieved by MA students corresponds to consolidated C1 for the degree in "Language sciences" and to C2 for the degree in "Translation and Interpretation" within the CEFR framework.

Sixty people entered the link on Labvanced (Finger et al. 2017), but two of them declared to have both deaf signing parents and were therefore redirected to an ending page, and three of them did not complete the task, so the final number of participants is 55.

Participants were divided into two groups depending on their frequency of use of LIS: 35 participants declared to use LIS either every day or several days a week and were coded as "frequent users" (Age: mean = 31, min = 22, max = 51. Mean years of LIS use = 11, min = 4, max = 31); 20 participants declared to use LIS some days a month or a year and were coded as "non-frequent users" (Age: mean = 26, min = 24, max = 35. Mean years of LIS use = 6, min = 4, max = 15).

Frequent and non-frequent users were unbalanced for years of exposure to LIS: Most non-frequent users had a maximum of 7 years of LIS exposure (N = 18) and a minimum of 4, whereas frequent users were more distributed, with 15 participants with a maximum of 7 years of LIS exposure and a minimum of 4 and 20 participants with at least 10 years of LIS exposure and a maximum of 31. For this reason, given our research questions, we decided to perform the analysis considering two different subsets of participants, so to be able to evaluate the effects of both dimensions of frequency of use and of years of use, one for each subset of participants.

Specifically, in the first set of analyses the sample consisted of 33 participants, frequent and non-frequent users, balanced for years of LIS exposure (from 4 to 7). With this sample, we were able to assess the effect of frequency of LIS use (see 3.1.1 and 3.2.1). In the second set of analyses, the sample consisted of 35 participants, only frequent users. With this sample, we were able to assess the effect of years of LIS use (see 3.1.2 and 3.2.2).

2.2 Materials and procedure

On Labvanced, we duplicated the LEX-I and SEN-I tasks used in Giustolisi et al. (2024). We invite the interested readers to look at the original work and its online materials.

As a summary, LEX-I is a semantic decision task (Is it edible?) and SEN-I a truth value judgment task (Is it true?). Items are lexical units in LEX-I and short sentences in SEN-I. In both tasks, materials consisted of videos of a CODA uttering the target item either in Italian (spoken language condition), in LIS (sign language condition), or in both languages simultaneously (code-blended condition).

The peculiarity of SEN-I is that the spoken and the signed sentences in the code-blended condition are not lexically aligned, as we selected items where the unmarked word order was different between Italian and LIS, as in the example (2) above.

2.3 Data analysis

Statistical analyses were performed using R version 4.3.2 R Core Team (2021). Firstly, we observed the overall accuracy with the purpose of removing any participant with an overall error rate higher than 20%. No participant was excluded following this criterion.

We measured accuracy, analyzed with generalized linear mixed models, and reaction times (RTs), analyzed with linear mixed models. RTs were calculated both with respect to the start of the video (RTs-video) and with respect to the start of the audio (RTs-audio). This means that we calculated two different RTs for the code-blended condition, which involved both video and audio, following the methodology adopted in Emmorey, Petrich & Gollan (2012) and Giustolisi et al. (2024), to which we refer for details. The comparison between the sign language condition and code-blended condition considered RTs-video, whereas the comparison between the spoken language condition and code-blended condition considered RTs-audio. RTs were analyzed with linear mixed models.

As mentioned in 2.1, we performed two sets of analysis: one compared frequent and non-frequent users balanced for years of exposure to LIS and one considered only frequent users comparing L2 LIS signers with a different time lapse of LIS exposure.

In the first set of analyses (frequent vs. non- frequent users), we considered as fixed factors Group (treatment coding with non-frequent LIS users coded as 0 and frequent LIS users coded as 1), Condition (treatment coding with code-blending coded as 0 and spoken language and sign language coded as 1) and their interaction.

In the second set of analyses (frequent users with different years of LIS use), we considered as fixed factors Group (treatment coding with non-frequent LIS users coded as 0 and frequent LIS users coded as 1), years of LIS use, and their interaction.

In all analyses, Likelihood ratio tests were performed to compare the model with the interaction with the model without the interaction. The interaction was retained if its inclusion determined a significant increase in the model's goodness-of-fit.

In all analyses, we hypothesized random intercepts for participants and items and by participants and by items random slopes for the effect of condition ((condition|participant) + (condition|item)). If the model with this random structure did not converge, we simplified the random structure omitting random slopes until we reached convergence and we always avoided random structures resulting in a singular fit.

3 Results

We first present the results of the lexical experiment (LEX-I) in Section 3.1 and then the results of the sentence experiment (SEN-I) in Section 3.2.

3.1 Lexical semantic decision task (LEX-I)

We begin by considering frequent and non-frequent LIS users balanced in years of LIS exposure. In this analysis, among conditions, we entered the model group (frequent vs. non-frequent users) as predictor (3.1.1). Then we considered frequent users with various years of exposure to LIS. In this second analysis, among conditions, we entered in the model the years of LIS use (mean centered) as predictor (3.1.2). In each section, we start analyzing accuracy and then RTs in the code-blended /

sign language comparison (RTs-video) and in the code-blended / spoken language comparison (RTs-audio).

3.1.1 Frequent vs. non-frequent LIS users balanced in years of LIS exposure (from 4 to 7)

A summary of accuracy data in the LEX-I task is given in Table 1. No significant Group*Condition interaction was found ($\chi^2(2) = 1.99$, $p = .37$). Accuracy in the sign-language condition was lower than in the code-blending condition and the difference was significant. Accuracy did not significantly differ between the spoken language and the code-blended condition. Frequent users were more accurate than non-frequent users, but again this difference was not significant (Table 2).

	Condition		
Group	Sign language M (SD)	Spoken language M (SD)	Code-blended M (SD)
Frequent users	89% (35%)	98% (13%)	98% (14%)
Non-frequent users	82% (39%)	98% (15%)	99% (12%)

Table 1: Overall mean accuracy for semantic decisions on LIS signs, Italian words, and LIS-Italian code-blends in frequent and non-frequent LIS users

Fixed factors	Estimate	SE	z-value	p-value
Condition: Sign Language	-2.04	0.67	-3.03	.002
Condition: Spoken Language	2.82	1.77	1.59	.11
Group: non-frequent users	-0.35	0.20	-1.74	.08

Table 2: Parameters of the model for accuracy in LEX-I (frequent vs. non-frequent LIS users)

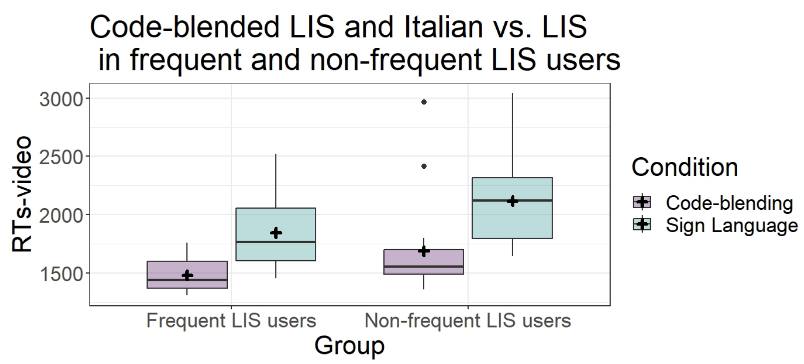

Figure 1: Box plots of by-subject mean RTs from the beginning of the video for semantic decisions to LIS/Italian code-blends and LIS signs in frequent and non-frequent LIS users. The + symbol corresponds to the mean value in each condition.

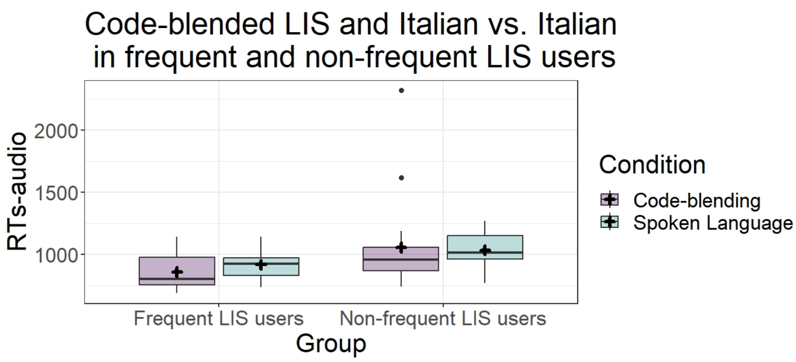

Figure 2: Box plots of by-subject mean RTs from the beginning of the audio for semantic decisions to LIS/Italian code-blends and Italian words in frequent and non-frequent LIS users. The + symbol corresponds to the mean value in each condition.

As for RTs, results are depicted in Figure 1 and Figure 2. Considering RTs-video, the Group*Condition interaction was not significant ($\chi^2(1) = 0.77$, $p = .38$). Participants were slower in the sign language condition compared to the code-blending condition and the difference was significant. Non-frequent users were in general slower than frequent users, and the difference was significant (Table 3). Also for RTs-audio the Group*Condition interaction was not significant ($\chi^2(1) = 1.04$, $p = .31$). RTs in the spoken language condition were comparable with that in the code-blending condition and the difference was not significant. Non-frequent users were in general slower than frequent users, and the difference was significant (Table 3).

Dependent variable	Fixed factors	Estimate	SE	t-value	p-value
RTs-video	Condition: Sign language	404.11	23.25	17.38	< .0001
	Group: Non-frequent users	233.86	101.25	2.31	.03
RTs-audio	Condition: Spoken language	13.66	41.60	0.33	.75
	Group: Non-frequent users	105.46	43.59	2.42	.02

Table 3: Parameters of the model for RTs in LEX-I (frequent vs. non-frequent LIS users)

3.1.2 Frequent LIS users with various years of LIS use (from 4 to 31)

Results on task accuracy are reported in Table 4. No significant interaction between condition and years of LIS use was found ($\chi^2(2) = 4.12$, $p = .13$). Participants were significantly less accurate in the sign language condition compared to the code-blending condition, whereas accuracy was comparable between the spoken language and the code-blending condition.

Condition		
Sign language	Spoken language	Code-blended
M (SD)	M (SD)	M (SD)
89% (31%)	98% (13%)	98% (12%)

Table 4: Overall mean accuracy for semantic decisions on LIS signs, Italian words, and LIS-Italian code-blends in frequent LIS users with 4 to 31 years of exposure to LIS

In general, accuracy increased as years of LIS exposure increased, however the trend did not reach significance (Table 5).

Fixed factors	Estimate	SE	z-value	p-value
Condition: Sign Language	-2.61	0.31	-8.41	$<.0001$
Condition: Spoken Language	-0.25	0.38	-0.65	.51
Years of LIS exposure	0.20	0.12	1.64	.10

Table 5: Parameters of the model for accuracy in LEX-I (frequent users, various years of LIS use)

Turning to RTs, results by condition are shown in Table 6. In both analyses, the condition by years of LIS use interaction was not significant (RTs-video: $\chi^2(1) = 3.58$, $p = .06$; RTs-audio: $\chi^2(1) = 0.17$, $p = .68$). Participants were significantly faster in the code-blending condition compared to the sign-language condition, whereas there was no significant difference between the code-blending condition and the spoken language condition. Considering RT-s audio, RTs significantly increased as years of LIS exposure increased (Table 7).

	Condition		
	Sign language	Spoken language	Code-blended
	M (SD)	M (SD)	M (SD)
RTs-video	1774 (586)		1564 (340)
RTs-audio		971 (219)	938 (336)

Table 6: Overall mean RTs-video in ms for semantic decisions on LIS signs and LIS-Italian code-blends and overall mean RTs-audio in ms for semantic decisions on Italian words and LIS-Italian code-blends in frequent LIS users with 4 to 31 years of exposure to LIS

Dependent variable	Fixed factors	Estimate	SE	t-value	p-value
RTs-video	Condition: Sign language	233.38	44.13	5.06	< .0001
	Years of LIS exposure	72.50	45.50	1.59	.12
RTs-audio	Condition: Spoken language	34.42	34.24	1.01	.32
	Years of LIS exposure	66.93	22.80	2.04	.006

Table 7: Parameters of the model for RTs in LEX-I (frequent users, various years of LIS use)

3.1.3 Summary of results in LEX-I

Similar to the CODA population studied in Giustolisi et al. (2024), we found an advantage for the code-blended condition over the sign language condition in L2 signers. The effect was found in both accuracy and RTs. We found an overall effect of group, with more frequent users being faster than non-frequent users. No effect of condition was found when comparing code-blending and spoken language, differently from what was observed with Italian CODAs.

3.2 Sentential truth value judgment task (SEN-I)

We now turn to reporting the results of the experiment involving full sentences in the three conditions. First, we present results comparing frequent and non-frequent users (3.2.1), and then the results assessing the effect of years of LIS use in frequent users (3.2.2). We start with the analysis of accuracy and then RTs for the code-blended / sign language comparison and then for the code-blended / spoken language comparison.

3.2.1 Frequent vs. non-frequent LIS users balanced in years of LIS exposure (from 4 to 7)

A summary of accuracy data in the SEN-I task is provided in Table 8. No significant Group*Condition interaction was found ($\chi^2(2) = 3.06$, $p = .22$). Accuracy in the sign-language condition was lower than in the code-blending condition and the difference was significant. Accuracy in the spoken language condition and in the code-blending condition was comparable. There was no significant difference in accuracy between frequent and non-frequent users (Table 9).

Group	Condition		
	Sign language M (SD)	Spoken language M (SD)	Code-blended M (SD)
Frequent users	77% (42%)	97% (17%)	98% (15%)
Non-frequent users	73% (45%)	98% (14%)	98% (15%)

Table 8: Overall mean accuracy for truth value judgments on LIS sentences, Italian sentences, and LIS/Italian code-blended sentences in frequent and non-frequent LIS users

Fixed factors	Estimate	SE	z-value	p-value
Condition: Sign Language	-2.94	0.23	-12.70	<.0001
Condition: Spoken Language	-0.05	0.30	-0.16	.87
Group: non-frequent users	-0.16	0.21	-0.78	.44

Table 9: Parameters of the model for accuracy in SEN-I (frequent vs. non-frequent LIS users)

As for RTs, results are depicted in Figures 3 and 4. Considering RTs-video, the Group*Condition interaction was not significant ($\chi^2(1) = 0.05$, $p = .82$). Participants were slower in the sign language condition compared to the code-blending condition and the difference was significant. RTs-video did not significantly differ between frequent and non-frequent LIS users (Table 10). For RTs-audio the Group*Condition interaction was not significant ($\chi^2(1) = 0.07$, $p = .79$). RTs in the spoken language condition were slightly faster than in the code-blending condition, but the difference was not significant. Just like for TRs-video, RTs-audio did not significantly differ between frequent and non-frequent LIS users (Table 10).

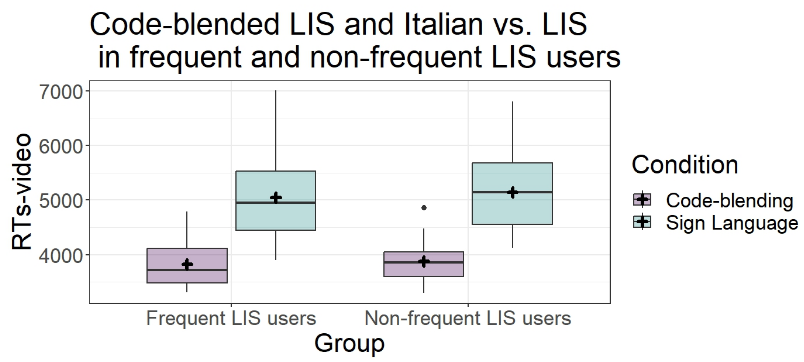

Figure 3: Box plots of by-subject mean RTs from the beginning of the video for truth value judgments on LIS-Italian code-blended sentences and LIS sentences in frequent and non-frequent LIS users. The + symbol corresponds to the mean value in each condition.

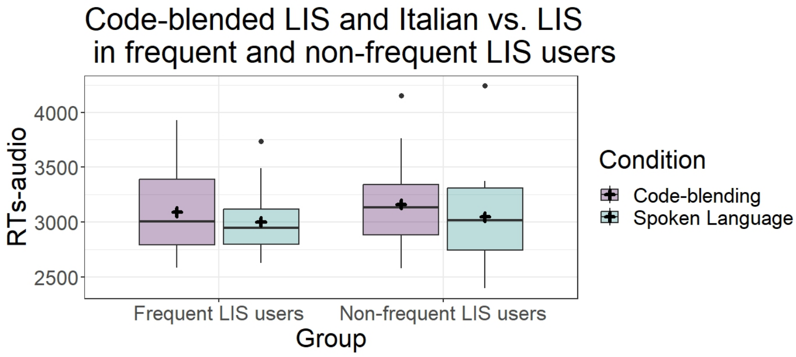

Figure 4: Box plots of by-subject mean RTs from the beginning of the audio for truth value judgments on LIS/Italian code-blended sentences and Italian sentences in frequent and non-frequent LIS users. The + symbol corresponds to the mean value in each condition.

Dependent variable	Fixed factors	Estimate	SE	t-value	p-value
RTs-video	Condition: Sign language	1339.26	128.30	10.44	< .0001
	Group: Non-frequent users	52.14	146.75	0.36	.73
RTs-audio	Condition: Spoken language	-95.64	53.20	-1.80	.08
	Group: Non-frequent users	49.48	127.39	0.39	.70

Table 10: Parameters of the model for RTs in SEN-I (frequent vs. non-frequent LIS users)

3.2.2 Frequent LIS users with varying years of LIS use (from 4 to 31)

Results on task accuracy are reported in Table 11. No significant interaction between condition and years of LIS use was found ($\chi^2(2) = 1.57$, p = .46). Parameters of the model without the interaction are reported in Table 12. Participants were significantly less accurate in the sign language condition compared to the code-blending condition, whereas accuracy was comparable between spoken language and code-blending. There was no significant difference in accuracy related to the years of LIS use.

Condition		
Sign language M (SD)	Spoken language M (SD)	Code-blended M (SD)
83% (38%)	98% (16%)	97% (16%)

Table 11: Overall mean accuracy for truth value judgments on LIS signs, Italian words, and LIS-Italian code-blends in frequent LIS users with 4 to 31 years of exposure to LIS

Fixed factors	Estimate	SE	z-value	p-value
Condition: Sign Language	-2.17	0.21	-10.31	<.0001
Condition: Spoken Language	0.05	0.27	0.18	.86
Years of LIS exposure	0.09	0.12	0.80	.43

Table 12: Parameters of the model for accuracy in SEN-I (frequent users, various years of LIS use)

RTs results are reported in Table 13. As for RTs-video, the condition by year of LIS use interaction was significant ($\chi^2(1) = 5.25$, $p = .02$). Basically, RTs were faster in the code-blending condition compared to the sign language condition, however this difference was modulated by years of LIS exposure (being maximal for those with few years of LIS exposure). Years of LIS exposure impacted RTs in the sign language conditions but not in the code-blending condition. Parameters of the model with the interaction are reported in Table 14, the interaction is depicted in Figure 5.

	Condition		
	Sign language M (SD)	Spoken language M (SD)	Code-blended M (SD)
RTs-video	4753 (1467)		3868 (781)
RTs-audio		3014 (595)	3225 (743)

Table 13: Overall mean RTs-video in ms for truth value judgments on LIS signs and LIS-Italian code-blends and overall mean RTs-audio in ms for truth value judgments on Italian words and LIS-Italian code-blends in frequent LIS users with 4 to 31 years of exposure to LIS

Considering RTs-audio, the condition by year of LIS use was not significant ($\chi^2(1) = 0.07$, $p = .80$). Regardless of the years of LIS use, participants were significantly faster in the spoken language condition compared to the code-blending condition.

Dependent variable	Fixed factors	Estimate	SE	t-value	p-value
RTs-video	Condition: Sign language	979.52	11.87	8.60	<.0001
	Years of LIS exposure	46.74	73.56	0.64	.53
	Condition*Years of LIS exposure	-234.12	100.78	-2.323	.026
RTs-audio	Condition: Spoken language	-220.35	64.88	-3.40	.001
	Years of LIS exposure	71.03	47.92	1.48	.15

Table 14: Parameters of the model for RTs in SEN-I (frequent users, different years of LIS use)

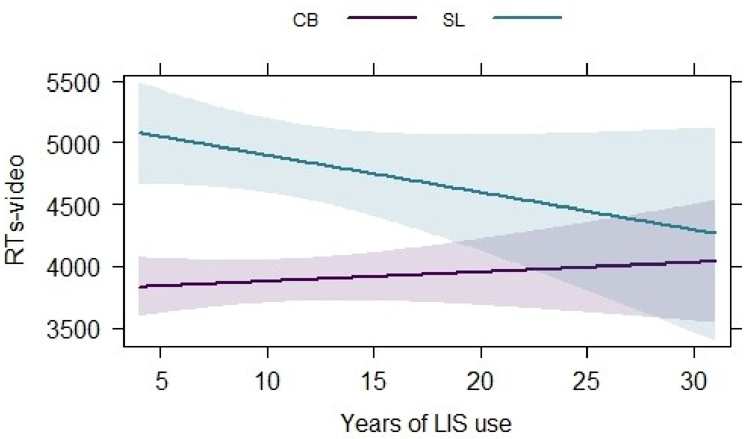

Figure 5: Effect plot of the years of LIS use by condition (CB = code-blended, SL = sign language) interaction found in the analysis of RTs-video of the truth value judgments on LIS-Italian code-blended sentences and LIS sentences.

3.2.3 Summary of results of the SEN-I test

Like in the case of the LEX-I test and like the Italian CODAs studied in Giustolisi et al. (2024), participants performed better in the code-blended condition than in the sign language condition in terms of accuracy and RTs, irrespectively of the frequency of use or years of exposure to LIS. Differently from what was found in the LEX-I test, no generalized effect of Group was found, neither for frequency of use nor for years of exposure. However, an interaction of years of exposure was found in the RT-video analysis (sign language vs. code-blending). Participants with less years of LIS use were slower in the sign language condition than in the code-blended condition when compared to participants with more years of LIS use. Finally, RTs-audio were slower in the code-blended condition compared to the spoken language condition in frequent LIS users, irrespectively of the years of LIS use.

4 Discussion

The aim of the present work was to verify if late bimodal bilinguals (hearing L2 LIS signers with Italian as L1) elaborate a bimodal input (Italian/LIS code-blending) differently from a unimodal one (either LIS only or Italian only) focusing on lexical access and on sentences comprehension. Participants varied as for frequency of LIS use and years of exposure to LIS. In our analyses, we first compared high and low frequent users of LIS relatively homogeneous in terms of years of exposure to LIS (from 4 to 7) and frequent users of LIS varying in years of exposure to LIS (from 4 to 32). As described in Section 2.1, participants to the present study were not evenly distributed for frequency of LIS use and years of LIS exposure. Specifically, the vast majority of non-frequent users had only between 4 and 7 years of LIS exposure, therefore we considered irrelevant to analyze the effects of years of LIS use in the present sample of non-frequent users, as there is a lack of participants who have been using LIS for a long time. Obviously, it would be of great interest to enlarge the sample of LIS L2 signers so as to have the possibility of considering long time non-frequent LIS users, an aspect we will investigate in future research. The baseline for comparison in the present analysis was the behavior of Italian CODAs as reported in Giustolisi et al. (2024), who show an advantage in processing bimodal lexical input compared to unimodal one, regardless of modality, but exhibit a cost in processing bimodal sentences compared to spoken sentences.

Before discussing the results of this study, it is important to point out two limitations of the present investigation. The first concerns the measure of frequency of use that we adopted, which is only based on one explicit question addressed to each participant and only quantified the frequency of use in terms of days a week/month/year, and not hours nor volume. The second limitation concerns LIS proficiency and involvement in the Deaf community, for which we have no measure. Having said that, and with all the caution that comes from these limitations, some facts appear to emerge that are worth discussing.

Overall, we found an advantage of code-blending over sign but not over speech. This is not surprising given that LIS is the L2 and Italian is the first language of all users. In fact, this is likely to be a dominance effect: we can assume that participants are faster when they have access to input in their first language. In the lexical task, non-frequent users were overall slower than frequent users. This general tendency can be explained by their competence in LIS, which is likely to be lower than in frequent users. However, non-frequent users were slower in general across the experiment, which might seem puzzling. Our speculation is that they were more cautious in performing a task that they globally perceived as more demanding than frequent users, being related to a language that they declared to use only rarely.

As for the advantage of code-blending with respect to spoken language in the lexical task that was identified as the crucial hallmark of CODA bilinguals, this did not emerge in our participants. This difference between CODAs and L2 hearing signers is prima facie interesting and might suggest that the advantage of lexical code-blending is out of reach for L2 users. But, given the limitations of our study mentioned above, and in particular the shallow measures of frequency of use that we adopted to determine our sample, it is not possible to raise conclusions beyond speculation from the null result reported in the present work. Notice, moreover, that our results are in contrast with what reported by Emmorey, Petrich & Gollan (2012), who found a code-blending advantage in the group they labelled as “late bilinguals”, which was composed of proficient L2 learners.

There are at least two differences between the two studies that might explain this contrast. One is the heterogeneity of the L2 populations included in the two studies: more than 60% of Emmorey, Petrich & Gollan (2012)’s participants were professional interpreters, and all participants declared to use ASL on a daily basis, while less than 50% of our participants were interpreters and frequent users declared to use LIS either on a daily or on a weekly basis, and it is thus possible that the American group was globally more proficient than the Italian one. It would be worth replicating the study in LIS/Italian with a population of L2 learners whose frequency of use and years of exposure are better measured and maximal (e.g., active interpreters with several years of activity and high involvement in the Deaf community) and verify whether the advantage of blending over the spoken condition arises in highly proficient late bilinguals.

Another difference between the two studies concerns the lexical items that were selected. Emmorey, Petrich & Gollan (2012) were able to control for the frequency of the English words included in the study. Interestingly, while the lexical advantage of code-blending was generalized for CODAs, it only emerged with frequent items in L2 learners. The Italian study did not control for this dimension, since the main criterion in Giustolisi et al. (2024) was that of selecting items that were likely to be accepted and understood across participants given the exceptional lexical variability that characterizes LIS on the Italian territory. As a result, lexical items of very diverging frequency were included. All in all, the data set included only a small number of high frequency items. In a follow-up study, it would be interesting to verify whether an advantage for blending arises in the presence of highly frequent lexical items.

As for the cause of the advantage of lexical bimodal stimuli, Emmorey, Petrich & Gollan (2012) attribute it to the redundant signals effect. Our results seem to indicate that only natives are able to profit from this effect independently from frequency. If this specificity of CODAs is confirmed, this might suggest that something more specific than a domain generic cognitive phenomenon such as the redundant signals effect might be at play in processing code-blending.

Considering sentences (SEN-I), the analysis of frequent and nonfrequent LIS users balanced in years of LIS exposure (4 to 7 years) yielded mainly null results, with the exception of slower RTs in the code-blending condition compared to the sign language condition, an effect that we attribute to a higher reliance on L1 compared to L2 in performing the task. The effect remains in frequent users only (4 to 32 years of LIS exposure).

Interestingly, frequent LIS users displayed an interaction between condition (sign language vs. code-blending) and years of LIS exposure, with faster RTs in the sign language condition for participants with more years of LIS exposure compared to those with less years of LIS exposure. This result might reflect the degree of competence in LIS as L2, that we expect to increase as years of LIS exposure increase.

Again, focusing on frequent users of LIS, irrespective of their years of exposure, the analysis shows slower RTs in the code-blended condition than in the Italian condition. This result replicates what Giustolisi et al. (2024) found for CODAs and it suggests that incongruent word orders generate a cost for processing code-blended sentential stimuli in all frequent LIS users, regardless of the age of first exposure to LIS.

In Giustolisi et al. (2024), the low performance of CODAs in the code-blended condition was explained by the cost of simultaneously parsing two divergent syntactic structures. We can speculate that a similar explanation applies here too, or that the cost is simply associated to linear misalignment. In order to better understand whether the source of the difficulty of incongruent blended sentences is the same for CODAs and L2 users, it would be interesting to turn to production data and verify whether CODAs and L2 users produce the same type and amount of incongruent code-blended sentences. We hope to perform such a study in the near future.

Acknowledgements

Materials, anonymized data, and scripts for the analysis are reported on the OSF page of the present project (osf.io/z8smh). Authors' contributions: CD, CG, BG and CB conceptualized the work. GM and CB took care of data collection. BG analyzed the data. All authors discussed the results. CD drafted the first version of the paper. CG, BG and CB revised the paper. All authors accepted the last version of the paper.

References

Baker, A. E. & B. Van Den Bogaerde. 2008. Codemixing in signs and words in input to and output from children. In C. Plaza-Pust & E. Morales-López (eds.), *Studies in Bilingualism*, vol. 38. Amsterdam: John Benjamins Publishing Company.

Bishop, M. & S. Hicks. 2005. Orange Eyes: Bimodal Bilingualism in Hearing Adults from Deaf Families. *Sign Language Studies* 5(2).
Branchini, C. & C. Donati. 2016. Assessing lexicalism through bimodal eyes. *Glossa: a journal of general linguistics* 1(1).
De Quadros, R. M. 2018. Bimodal Bilingual Heritage Signers: A Balancing Act of Languages and Modalities. *Sign Language Studies* 18(3).
Donati, C. & C. Branchini. 2013. Challenging linearization: Simultaneous mixing in the production of bimodal bilinguals. In *Challenges to Linearization.*
Emmorey, K., J. A. F. Petrich & T. H. Gollan. 2012. Bilingual processing of ASL–English codeblends: The consequences of accessing two lexical representations simultaneously. *Journal of Memory and Language* 67(1).
Emmorey, K., J. A. F. Petrich & T. H. Gollan. 2013. Bimodal Bilingualism and the Frequency-Lag Hypothesis. *Journal of Deaf Studies and Deaf Education* 18.
Finger, H., C. Goeke, D. Diekamp, K. Standvoß & P. König. 2017. Labvanced: a unified javascript framework for online studies. In *International conference on computational social science (ic2s2)*.
Giustolisi, B., A. Jaber, C. Branchini, C. Geraci & C. Donati. 2024. Processing code-blending beyond the lexical level: evidence for a double syntactic derivation? *Bilingualism: Language and Cognition* 27(5).
Kanto, L., K. Huttunen & M.L. Laakso. 2013. Relationship Between the Linguistic Environments and Early Bilingual Language Development of Hearing Children in Deaf-parented Families. *Journal of Deaf Studies and Deaf Education* 18(2).
Miller, J. 1986. Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics* 40(5).
Petitto, L. A., M. Katerelos, B. G. Levy, K. Gauna, K. Tétreault & V. Ferraro. 2001. Bilingual signed and spoken language acquisition from birth: implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language* 28(2).
R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. `https://www.R-project.org/`.