Bilingualism: Language and Cognition

cambridge.org/bil

Processing code-blending beyond the lexical level: evidence for a double syntactic derivation?

Beatrice Giustolisi[1], Angélique Jaber[2], Chiara Branchini[3], Carlo Geraci[4] and Caterina Donati[2]

[1]Department of Psychology, University of Milano-Bicocca, Milan, Italy; [2]University Paris Cité, CNRS, Laboratoire de Linguistique Formelle, Paris, France; [3]Department of Linguistics and Comparative Cultural Studies, Ca' Foscari University of Venice, Venice, Italy and [4]Institut de Jean Nico, PSL University, CNRS, Paris, France

Research Article

Cite this article: Giustolisi, B., Jaber, A., Branchini, C., Geraci, C., & Donati, C. (2024). Processing code-blending beyond the lexical level: evidence for a double syntactic derivation? *Bilingualism: Language and Cognition*, **27**, 927–938. https://doi.org/10.1017/S1366728923000901

Received: 31 January 2023
Revised: 23 November 2023
Accepted: 27 November 2023
First published online: 16 January 2024

Keywords:
bimodal bilingualism; code-blending; CODA; co-activation; semantic judgement task; truth value judgement task

Corresponding author:
Beatrice Giustolisi;
E-mail: beatrice.giustolisi@unimib.it

Abstract

Bimodal bilinguals master languages in two modalities, spoken and signed, and can use them simultaneously due to the independence of the articulators. This behavior, named code-blending, is one of the hallmarks of bimodal bilingualism. Lexical experiments on production and comprehension in American Sign Language/English showed that blending is not cognitively costly and facilitates lexical access. In this work, we replicated the blending advantage in lexical comprehension for hearing bimodal bilinguals with two other language pairs, French Sign Language (LSF)–French and Italian Sign Language (LIS)–Italian, and we explored whether the facilitation is also found at the sentential level. Results show that blended utterances for languages with incongruent word order like LIS–Italian were processed slower than monolingual utterances, while no difference was found when the word orders were congruent (LSF–French). We discuss these findings in light of linguistic theories of syntactic structure derivation in bimodal bilinguals.

1. Introduction

Sometimes, doing two things together is easier than doing only one.

An outstanding and fascinating fact about bilingualism that has emerged over and over is that the two languages a bilingual person knows are always active to some degree when they are engaged in tasks of various types, even those explicitly requiring one language only. This has been found in reading tasks (Dijkstra & Kroll, 2005), in listening tasks (Marian & Spivey, 2003), in spontaneous speech (Kroll et al., 2006), and in a family of Stroop tasks (Giezen et al., 2015). Cross-language activation has been shown to be effective at various levels of processing. For instance, cross-linguistic priming effects have been found at both lexical (Finkbeiner et al., 2004) and syntactic (Loebell & Bock, 2003) levels. These effects are not only restricted to languages sharing a similar writing or phonological system: they were documented in English–Japanese bilinguals (Ikeda, 1998), and even across modalities in bimodal sign-speech bilinguals (Morford et al., 2011; Ormel et al., 2012; Shook & Marian, 2012). In sum, there is growing consensus that bilinguals do not "unplug" the language that they are not using, even when it would be beneficial to do so.

Within this framework, the case of bimodal bilingualism is the focus of our paper. Bimodal bilinguals are competent users of a sign language and a spoken language, and thus have access to two largely independent articulatory channels: the visual-manual channel and the auditory-vocal channel. In this setting, no articulatory constraint prevents bimodal bilinguals from using both languages simultaneously, which makes them an ideal test bed for investigating the interaction of the two languages and the limits of their co-activation BEYOND articulatory constraints. The possibility of articulating two languages simultaneously unveils important questions about the nature of their representations and the possible costs or benefits associated with accessing two representations at the same time.

Among bimodal bilinguals, we shall only focus on hearing bilinguals in this paper, and in particular, on populations of CODAs (Children Of Deaf Adults), hearing adults who acquired spontaneously the sign language in their family and the spoken language in the larger surrounding environment.

The paper is organized as follows: section 1.1 provides an overview of the specificities of language mixing in bimodal bilinguals' production, and section 1.2. summarizes the few experimental results available on their comprehension, which is the focus of the present

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

This article has earned badges for transparent research practices: Open Data and Open Materials. For details see the Data Availability Statement.

investigation. Section 2 introduces the present study, outlining the main hypotheses at stake. Section 3 presents the experimental methods; section 4 illustrates their results. Section 5 is a discussion, and section 6 concludes the paper.

1.1. Code-blending in bimodal bilinguals: production

The possibility of simultaneous two-language use is fully exploited in bimodal bilinguals. Differently from unimodal bilinguals, they tend not to switch from one language to the other during the same utterance, sentence, or phrase (Muysken, 2000). Rather, they prefer to CODE-BLEND (Emmorey et al., 2008), meaning that they sometimes produce fragments belonging to the two languages simultaneously. This appears to be true for all the bimodal populations that have been observed so far, be they adults or children (American Sign Language (ASL)–English in Bishop & Hicks, 2005; Dutch Sign Language (NGT)–Dutch in Baker & Van den Bogaerde, 2008; Quebec Sign Language (LSQ)–French in Petitto et al., 2001; Brazilian Sign Language (LIBRAS)–Portuguese in de Quadros, 2018; Italian Sign Language (LIS)–Italian in Bishop et al., 2007, Donati & Branchini, 2013, Branchini & Donati, 2016; Finnish Sign Language (FinSL)–Finnish in Kanto et al., 2013).

For example, Emmorey et al. (2008) analyzed the mixed utterances produced by a group of American adult CODAs in spontaneous conversation and narrative tasks, and found that code-blending accounted for 36% of the overall production and 98% of the mixing data. Similar patterns have been reported for KODAS (i.e., Kids Of Deaf Adults, a population of hearing bimodal bilingual children: NGT/Dutch – Baker & Van den Bogaerde, 2008; LSQ/French – Petitto et al., 2001). While the vast majority of studies focuses on HEARING bimodal bilinguals, and so do we in this paper, this pattern appears to extend to Deaf children and adult bilinguals with oral training, at least according to the few studies available (e.g., Fung & Tang, 2016; Rinaldi & Caselli, 2014; Rinaldi et al., 2021). In a group of adult Deaf bilinguals of Hong Kong Sign Language (HKSL)–Cantonese observed by Fung and Tang (2016), code-blends accounted for around 40% of the overall production, and above 95% of mixing data.

Going back to hearing bilinguals, in the vast majority of cases reported in the literature, code-blended productions are syntactically and semantically congruent – hence exhibiting ALIGNED and SEMANTICALLY EQUIVALENT signed and spoken content: 81% for Emmorey et al. (2008); more than 80% in the Dutch study; 75% in the Canadian one (Petitto et al., 2001). An example of such syntactic/semantic congruent blends where this alignment extends to the entire sentence is given in (1), taken from Emmorey et al. (2008, p. 48).

(1) Eng: I don't think he would really live *Full congruent blend*
ASL: PRO NOT THINK REALLY LIVE[1]

In (1) two full semantically equivalent utterances are produced simultaneously and all their components are syntactically aligned.

As an important clarification, congruent blends are rarely fully so due to some fundamental differences between sign languages and spoken languages grammar. A number of functional elements that can typically surface in spoken languages, such as auxiliaries, prepositions, articles, clitics, and conjunctions, do not exist in sign languages as documented so far, and are therefore never «paired» in sign when uttered in a code-blend. The same holds, on the other hand, for so-called classifiers in sign languages, i.e., « morphemes with a non-specific meaning, expressed by particular configurations of the hands and which represent entities by denoting salient characteristics » (Zwitzerlood, 2012, p. 158), which have no grammatical equivalent in spoken languages (see de Quadros et al. (2020) for an interesting analysis of what happens when blending involve depicting classifiers).

Interestingly, congruent code-blends can be produced in one of two different ways: the congruency and systematic alignment can be due to the fact that the two languages happen to prescribe the same word order which is thus consistent across the two strings. This is the case we shall see with LSF–French in the experiments described below, and it is further illustrated in (2) with an ASL–English blend. In the other case, congruency and alignment can be obtained by imposing the word order corresponding to one language only to both strings with a process akin to syntactic calque (illustrated in 3)[2].

(2) Eng: You're not getting one *Consistent congruent blend (no calque)*
ASL: PRO NOT GET ONE (Bishop, 2011, p. 228)

Each language string in (2) exhibits a full-fledged utterance that would be acceptable even uttered alone, in a unimodal context. This is very different from what is observed in (3).

(3) a. Engl: Cute plus deaf poor lalalala *ASL-based congruent blend with calque*
ASL: CUTE PLUS DEAF POOR PITY
'(She's) cute, deaf, and poor. What a pity'. (Bishop, 2011, p. 226)
b. It: Una bambina va allo zoo *Italian-based congruent blend with calque*
a girl goes to.the zoo
LIS: GIRL GO ZOO (Branchini & Donati, 2016, p. 11)
'The girl goes to the zoo.'

Here, both in (3a) and (3b) only one of the two language strings makes up a grammatical sentence that could be produced in a monolingual context, while the other string is a syntactic calque of the construction prescribed by what is often called the matrix language: ASL in (3a) and Italian in (3b).

Be that as it may, in most cases of spontaneous production, only some portions of the utterance are bimodal and congruently aligned, and only one language is fully responsible for the utterance content/word order. These PARTIAL blends are illustrated in (4), taken from Bishop (2011, p. 225): (4a) is an English-dominant code-blend, while (4b) is an ASL-dominant one.

(4) a. Engl: I think it is mostly English *English-dominant partial blend*
ASL: MOST ENGLISH
'I think it is mostly English'
b. Engl: but I knew if we did not *ASL-dominant partial blend*
ASL: BUT IX-1 KNOW IF WE NOT, LIFEZERO#L-I-F-E
'But I knew if we did not, we'd have no life.'

Turning to incongruent code-blends, where some mismatch occurs between the two language strings, some further distinctions need to be made, based on the empirically available cases. On the one hand, a code-blend can be SEMANTICALLY incongruent in one of two very different ways. First, it can be incongruent because the two language fragments express a slightly different content, as in (5).

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press

(5) It: Le meduse non c'erano *Semantically incongruent blend*
the.PL jellyfish.PL NEG there.be.PST.3PL
LIS: JELLYFISH SEE NOT
It: 'The jellyfish were not there.' / LIS: 'I didn't see the jellyfishes.'
(adapted from Branchini & Donati, 2016, p. 18)

In this interesting code-blend, observed in the spontaneous conversation of a 10 year old LIS–Italian KODA, the blend is incongruent because it contains two simultaneous sentences that have a slightly different meaning (*there were no jellyfish/ I saw no jellyfish*). In general, semantically incongruent code-blends of this kind are at best anecdotal according to the studies cited above. Semantic incongruence is typically more limited, with one of the two languages simply providing a piece of more specific information (*Twitty/bird* in (6)), but still compatible with a redundant description of the same event in the two blended languages.

(6) Eng: Tweety has binoculars too *Semantically incongruent blend*
ASL: BIRD HAVE BINOCULARS (Emmorey et al., 2008, p. 49)
'Tweety/the bird has binoculars too.'

But code-blend can be semantically incongruent in a completely different sense, when the fragments simultaneously provided by the two languages are compositionally integrated to form a unique proposition. An example is given in (7).

(7) Can: bei. *Integrated semantically incongruent blend*[3]
give
HKSL: BISCUITS.
biscuits
'Give (me) biscuits.' (CC 3;0.13 Fung & Tang, 2016, p. 768)

This second type of semantically incongruent code-blend is slightly more frequent in existing corpora.

On the other hand, code-blends can be SYNTACTICALLY incongruent – hence displaying the same lexical items in the two languages but not aligned into the same word order. This can typically occur when the two languages display different word orders, as in LIS–Italian. An example is given in (8).

(8) It: Non ho capito *Syntactically incongruent blend*
NEG have.1s understand.Pst
LIS: UNDERSTAND NOT
'I didn't understand.' (Donati & Branchini, 2013, p. 108)

However, most pairs of languages involved in existing studies display similar word orders, at least as far as lexical items are concerned. So, it is possible that the predominance of syntactically congruent code-blends that is observed in those studies is in part due to the nature of the sample, which is biased in favor of congruent code-blends because the grammars of the paired languages have similar word orders[4]. In fact, works on code-blends in two syntactically distant languages like LIS and Italian found several cases of syntactically incongruent blends (Branchini & Donati, 2016; Donati & Branchini, 2013). Even if Italian and LIS are rather free word order languages, the canonical word order in an out-of-the-blue context is divergent, with LIS preferentially head final (hence SOV, to say it simply: Branchini & Mantovan, 2020) and Italian consistently head initial (hence SVO). Although no quantitative measure is provided in those studies, Donati and Branchini observed a consistent production of syntactically incongruent code-blends that were semantically congruent. These data emerged in a variety of contexts, like spontaneous conversations, narrative tasks, and repetition tasks (see also Lillo-Martin et al., 2016 for a discussion). Crucially, these productions are congruent from the point of view of lexical access but misaligned with respect to the order of the constituents in the two languages. Their status should be further investigated, a point to which we shall shortly turn.

Summarizing, code-blending is very natural in the spontaneous production of bimodal bilingual CODAs, and it comes in various flavors. Blending can be full (when two complete utterances are produced simultaneously) or partial; syntactically and semantically congruent or incongruent. These distinctions are summarized in the graph below (Figure 1).

Going back to a more general level, as discussed by Emmorey et al. (2008), the fact that bimodal bilinguals produce more code-blending than code-switching is *per se* a confirmation that double selection is less costly for bilinguals than inhibition, at least at the lexical level. Otherwise, we would expect bimodal bilinguals to prefer code-switching over code-blending. Magnetoencephalography data have indeed confirmed that dominant language inhibition requires more cognitive resources than simultaneous lexical selection and production in both languages of bimodal bilinguals (Blanco-Elorrieta et al., 2018).

Related to this, Emmorey et al. (2012) investigated the response latencies of a group of American CODAs (N = 18) and a group of hearing late learners of ASL (N = 20) in a picture-naming task where participants were asked to provide ASL signs only, English words only or ASL–English congruent blends. The authors found that blending did not slow lexical retrieval for ASL and actually facilitated access to low-frequency signs. Blending delayed speech production because bimodal bilinguals synchronized English and ASL lexical onsets.

1.2. Code-blending in hearing bimodal bilinguals: comprehension

Turning to comprehension, if dual lexical access is less costly than inhibition, we should also expect congruent blends to be processed more easily (hence quicker) than unimodal items. In the just cited study (Emmorey et al., 2012), Emmorey and collaborators tested the comprehension of code-blended lexical items in 18 American CODAs and 25 hearing late learners of ASL. They used a semantic categorization task (*Is the item edible*?) and compared semantic decision latencies for ASL-English congruent lexical blends with those for ASL signs alone and audiovisual English words alone.

They did not directly compare RTs for ASL and English because possible confounding effects of language modality were present in the stimuli (i.e., visual-manual signs vs. auditory-vocal words). Thus, for each language, they compared RTs and accuracy in the unimodal condition and in the blended one. The results showed that code-blending speeded up responses with respect to both languages, clearly showing that L1/L2 simultaneous processing does not require additional cognitive resources compared to the processing of single L1 or L2 expressions.

To the best of our knowledge, no study has yet experimentally investigated the comprehension of SENTENTIAL code-blends, where lexical items need not only to be accessed but also to be INTEGRATED syntactically, or compared the comprehension of syntactically congruent and incongruent blends.

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press

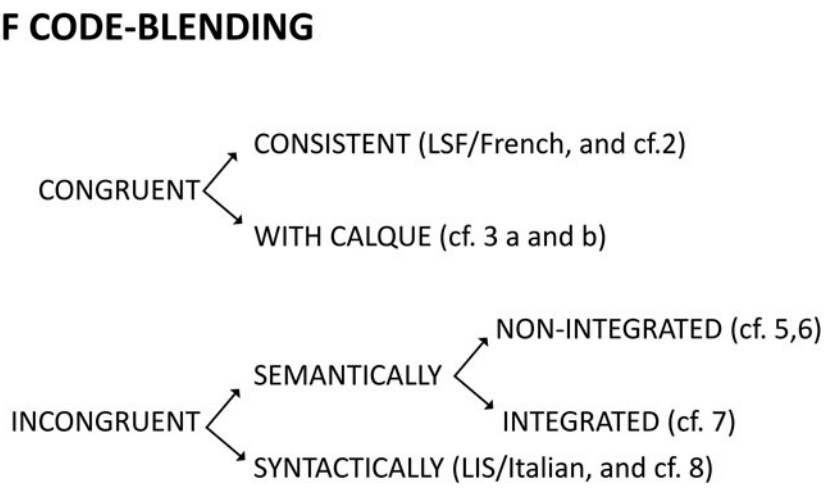


Figure 1. A schematic summary of the types of congruent and incongruent code-blendings

2. The present study: comprehending lexical and sentential code-blendings

The first aim of the present study is to replicate Emmorey et al. (2012)'s original finding on the comprehension of code-blended lexical items with new populations of CODAs: LSF–French bilinguals and LIS–Italian bilinguals. We thus designed two fully parallel online experiments (LEX-F, for the French population, and LEX-I, for the Italian one) involving lexical items under three conditions: spoken language (French; Italian), sign language (LSF; LIS) and code-blending (LSF– French; LIS– Italian), asking participants to respond as soon as possible to a simple semantic categorization question (*Is it edible?*). The hypothesis is that if the advantage found by the original study is a consistent feature of bimodal bilingualism, we should be able to replicate the same results in two new language pairs.

But, as we have reported in the examples above, the types of code-blends that bimodal bilinguals spontaneously produce are not limited to isolated words, they can rather involve entire sentences or, at least, constituents. The second aim of this study is to test whether the advantage of blending observed in Emmorey et al. (2012) also holds when full sentential blends are involved.

We believe this question is of general importance for the study of bilingualism because the advantage observed by Emmorey and her group for blended lexical items can be seen as yet another reflex of language co-activation: given that lexical co-activation is costless for bilinguals, accessing two lexical entries in processing lexical blends is faster than suppressing one lexical entry in processing unimodal stimuli. But, as we mentioned earlier, most studies on language co-activation and its facility focus on pure lexical co-activation. A natural question that arises is: do bilinguals only co-activate lexical items, or do they extend this co-activation to syntactically organized items? Studying whether code-blended sentences provide a processing advantage over unimodal sentences might contribute important evidence for addressing this question. However, since processing sentential structures is notoriously more complex than single-word processing, co-activation at the syntactic level might also translate into a processing disadvantage.

As a terminological clarification, this question is meaningful if considered according to mainstream models of the syntax-lexicon interface (so-called lexicalism, stemming from Chomsky, 1970). Within this view, syntax is feature-driven, in the sense that syntactic derivations read and obey the information that is encoded in lexical items. When bilinguals co-activate two lexical items in their two languages, do they also build two syntactic representations stemming from the features encoded in the words? Other "constructivist" models (e.g., Distributed Morphology, stemming from Halle & Marantz, 1993), are even more permeated by the feature-driven approach up to the point that the building of syntactic structures is treated as a pre-requisite to lexical insertion. Under this view, the question would be: does co-activation work at the level of lexical insertion only, or does it proceed directly from features?

We designed two similar extensions of the original lexical comprehension experiment (SEN-F and SEN-I), this time including full sentences, again presented under three conditions: either spoken language (French; Italian), or sign language (LSF; LIS), or code-blending (LSF–French; LIS–Italian), asking participants to answer as soon as possible a simple truth value question (*Is the sentence generally true?*). Crucially, all the items were constructed in such a way that the compositional integration of the verb with each of its arguments was necessary to provide the required truth judgment. An illustration is given in (9).

(9) example from SEN-F

a.		Les lions	mangent	les croissants	*Spoken French condition*
		The lions	eat	the croissants	
b.		LION	EAT	CROISSANT	*LSF condition*
		lion	eat	croissant	
c.	Fr:	Les lions	mangent	les croissants	*Code-blended (LSF–French) condition*
	LSF:	LION	EAT	CROISSANT	

'Lions eat croissants.'

In order to evaluate the truth of this sentence, whatever its condition of presentation, it is not enough to simply juxtapose the lexical meaning of the three words, but it is rather necessary to integrate them into a syntactic structure, by which 'lions' is the agent of the eating event, whose theme is 'croissants' (There is nothing wrong with the association of *lions* and *eating*, or with that of *eating* and *croissant*; still the sentences in (9) are generally false because lions do not typically eat croissant).

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press

This is where the replication of the same design in two different language pairs becomes of special importance. French and LSF are both head initial languages and can exhibit the same word order being at the same time consistent with their grammar. Therefore, all the sentential code-blends used in the SEN-F experiment are syntactically and semantically congruent, as in (9). Italian and LIS are typologically opposite, and therefore the sentential code-blends of the SEN-I experiment are all syntactically (but, crucially, not semantically) incongruent. To illustrate, the equivalent of (9) in SEN-I corresponds to (10), where the blended condition (10c) is syntactically incongruent.

(10) example from SEN-I

a.		I leoni	mangiano	le brioches	*Spoken Italian condition*
		The lions	eat	the croissants	
b.		LION	CROISSANT	EAT	*LIS condition*
		Lion	croissant	eat	
c.	It.	I leoni	mangiano	le brioche	*Code-blended (LIS–Italian) condition*
	LIS	LION	CROISSANT	EAT	

'Lions eat croissants.'

Here too, no matter the condition of the presentation, the only way to compute and assess the truth of the sentence is to syntactically integrate all of its components.

With this new pair of experiments, as we said, the first question we wanted to address was whether the same advantage of code-blending that Emmorey and her group found in the comprehension of single lexical items could be observed in a more complex task involving a full sentence. Concerning the impact of congruent and incongruent code-blends in this task, the hypotheses at stake are intricate because several factors may bias the outcome in various directions.

First, there might be a processing reflex of the advantage of double lexical access in the code-blend condition, which should result in a facilitation bias when the information is aligned, as already documented for categorization tasks in Emmorey et al. (2012), and might result in a penalization if the information is not aligned. Following this line, a facilitation effect is expected in the code-blended condition over the unimodal condition(s) in the congruent experiment (SEN-F), because lexical alignment across modality would speed processing (hence higher accuracy and shorter RTs). In the incongruent experiment (SEN-I), we expect a negative effect on the code-blended condition because lexical information coming from the two modalities is misaligned.

Second, there could be an effect induced by the temporality of access to the information needed to compute the truth value of the sentence. For example, the truth value of sentences with incongruent blends like (10c) can be computed earlier than their unimodal counterpart. In fact, the temporal alignment of 'LION/leoni – CROISSANT/eat' in (10c) is already enough to determine whether the blended sentence is true or not, without waiting the end of the utterance. Such an effect is not expected in sentences with congruent blends like (9c) because the systematic alignment of each cross-modal element does not allow for an early access to the truth value of the sentence. Following this line, an advantage for the code-blended condition over the unimodal condition(s) is expected in the incongruent experiment (SEN-I) and not in the congruent experiment (SEN-F).

However, sentences are not just sequences of words/signs: they are associated with a syntactic structure, which also needs to be processed in order to perform the task of the experiments we designed. In this respect, there is a debate on the syntax of code-blended sentences displaying two syntactically congruent (cf. 2 and 9) and incongruent (cf. 8 and 10) strings. Lillo-Martin et al. (2016) claim that blended sentences always involve only one shared structure that is lexicalized twice, either with the same linear order (as in congruent blends) or with two different ones (through two different linearization algorithms), as in incongruent blends. This proposal is based on the observation that in most production corpora the majority of code-blends are congruent, and that in many cases this happens because the word order of one language is imposed onto the blended language string, in what looks like a syntactic calque (see (3) above). Branchini and Donati (2016) maintain that code-blended sentences can also be associated with two syntactic structures, one for each utterance, when both strings comply with the word order required by each language and calque is not at stake. These structures can be either similar (as in non-calque congruent blends *à la* LSF–French: (9)) or different (as in incongruent ones *à la* LIS–Italian: (10)). This alternative proposal is based on the observation that blends that involve no calque and are thus either congruent blends, like LSF–French, or syntactically incongruent, like LIS–Italian, display morphosyntactic properties that are strongly divergent from what is visible in calqued blends, and call for the activation of two distinct grammars, beyond simple linearization[5].

According to Lillo-Martin et al.'s proposal, congruent blends have one structure which is linearized once, while incongruent blends have one structure which is linearized twice. Following this proposal, no bias is expected to favor or disfavor the blended condition over the unimodal condition(s) in the congruent experiment (SEN-F) because in all conditions there is just one structure to be processed. In the incongruent experiment (SEN-I), the expectation is that of a negative bias on the code-blended condition due to the costs of connecting two strings with different linear orders to a single structure.

According to Branchini and Donati's proposal, incongruent blends and non-calque congruent blends like the ones used in the stimuli of our experiments always involve two separate structures, therefore the expectation is that blended expressions are always penalized over non-blended expressions no matter whether congruent (SEN-F) or incongruent (SEN-I). The size of this penalization might be modulated by the divergence of the two syntactic structures to be processed. So, a larger negative bias is expected in SEN-I with respect to SEN-F.

Clearly, each factor might operate on an independent dimension, and a quantification of the strength of the biases is complicated to determine. Nonetheless, one can observe that the effect of the factors is predicted to be mostly negative for incongruent blends (penalization in the processing of the code-blended condition over the unimodal condition in SEN-I), where only early access to meaning might give a positive bias.

The picture is more differentiated for congruent blends: a processing facilitation is expected if one structure and one linearization is involved due to the positive effect of lexical alignment; while if two separate syntactic structures are involved, the positive effect of lexical alignment might be overshadowed by the extra costs imposed by syntax. Taking all these aspects into consideration, two alternative predictions can be formulated: Under one account, the combined effect of Lillo-Martin et al.'s proposal and the effects of lexical alignment should lead to facilitation in

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press

the processing of the code-blended condition over the unimodal condition in the congruent experiment (SEN-F).

Under the other account, the predictions for the congruent experiment (SEN-F) are more nuanced and crucially depend on the strength of the positive bias exerted by lexical alignment and the negative bias exerted by computing two structures (the Branchini and Donati's proposal) that are nevertheless very similar. Thus, if the positive bias and the negative bias have equal strength, the expectation is to find no difference between the bimodal and the unimodal conditions. If the positive bias is stronger than the negative bias, the expectation is to find a facilitation for the code-blended condition (in this particular case the two hypotheses make the same predictions), while if the negative bias is stronger than the positive one, the opposite prediction is expected.

3. Methods

The four experiments we conducted in this study were similar in methods and involved the same participants in each country: more precisely, the same group of participants in France took LEX-F and SEN-F, while the same group of participants in Italy took LEX-I and SEN-I. Within this experimental setup, results of the lexical experiment constitute a crucial baseline against which to interpret the sentence task results.

Due to the COVID-19 restrictions that have made the access to the peculiar population of hearing native bimodal bilinguals even more difficult, data collection was web-based. This is another reason why the replication of the original lexical experiment of Emmorey and colleagues was a necessary step before moving to the study of the effects of code-blending in sentence comprehension.

All materials are stored on the OSF repository of the present paper (osf.io/z8smh/).

3.1 Participants

Participants were recruited through flyers on social media and snowball sampling in both countries. The link with the experiment was sent via e-mail to 72 CODAs (the criteria of inclusion were to have at least one deaf signing parent and to self-define as bilingual), 40 from France (LSF-French bimodal bilinguals), and 32 from Italy (LIS-Italian bimodal bilinguals). However, several of them did not complete the experiment, therefore the final sample consisted of 25 French CODAs (20 females, mean age = 34 yrs, sd = 11yrs) and 25 Italian CODAs (17 females, mean age = 37 yrs, sd = 10 yrs)[6].

Before taking part to the experiments, participants answered a brief biographical questionnaire. As for French CODAs, 16 of them (64%) declared using French and LSF daily, 6 of them (24%) French on a daily basis and LSF several times a week, and the remaining 3 (12%) French on a daily basis and LSF several times a month. 7 French CODAs (28%) had high school diploma as their highest qualification, 5 (20%) a bachelor degree, 10 (40%) a master degree, 1 (4%) a doctoral degree, and 2 selected "*Autre*" (i.e., none of the proposed choices) in the questionnaire.

Considering Italian CODAs, 21 of them (84%) declared using Italian and LIS daily, whereas 2 of them (8%) declared using Italian every day and LIS several times a week, and other 2 (8%) Italian every day and LIS several times a month. 11 Italian CODAs (44%) had high school diploma as highest qualification, 8 (32%) a bachelor degree, 5 (20%) a master degree and 1 (4%) a doctoral degree.

All participants received 10€ for their participation, as an Amazon Gift Card.

The study was approved by the Ethics Committee of the Department of Psychology at the University of Milano-Bicocca. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

3.2. Lexical task: materials

For the two language pairs, materials consisted of a list of 90 concrete nouns that could refer either to an edible/drinkable entity (e.g., *cheese, pasta, juice*; N = 45) or to a non-edible/non-drinkable entity (e.g., *blanket, t-shirt, pencil*; N = 45). As selection criteria, we considered only nouns for which the corresponding sign was not a compound, to maximize sign-word alignment in code-blends. Moreover, we selected nouns for which the corresponding sign does not require an obligatory oral component[7], which would be lost in the code-blended condition. We also selected only signs that can have only one meaning when translated into the corresponding sign language. Last but not least, since lexical variation is widespread among sign languages, we selected common variants trying to avoid specific regional signs. Even if the vast majority of items between LEX-F and LEX-I corresponds to the same entity, our criteria made it impossible to have exactly the same list in the two experiments.

Two actors (a LSF-French CODA and a LIS-Italian CODA) produced the 90 target items plus 4 additional training items in three different conditions: spoken, signed, and code-blended (spoken + signed). Each CODA actor was filmed while producing the various items, maintaining a constant framing in the three modalities, with a frequency of 25 frames per second. For the spoken language condition, the two actors were instructed to keep their hands one above the other beneath their waist while pronouncing the words. The same position was used as starting and final position for each item in the sign language and in the code-blending condition. For the sign language condition, we asked our actors not to produce any mouth movement. We are aware that this is not an ecological choice, since signs are often accompanied by different kinds of movements of the mouths – however, since some of them have a direct link with spoken language, this option was the only one assuring that in the sign language condition, there wasn't any kind of spoken language input (for a thorough discussion see again footnote 7). Then, the video was edited to have one clip for each item in each condition.

The same person edited both French and Italian videos using Apple iMovie according to the following criteria: for items in the spoken condition, each clip was cut 9 frames (i.e., 360 ms) before and after the frames containing the actual sound of the word; for items in the signed and in the bimodal conditions, each clip began with the frame in which the hands began to move from the resting position and ended with the frame in which the hands were back still in the resting position.

For each language pair, the 90 target items were divided into three lists of 30 items each, 15 edible and 15 non-edible. For each language pair, the three lists were balanced for spoken language frequency, based on film subtitles corpora (for French – New et al., 2007; for Italian – Crepaldi et al., 2015). Currently, there are no frequency databases for LSF nor LIS.

In their relevant language pair and for each condition, each participant received one list, with lists counterbalanced across

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press

participants, so that all items appeared in all conditions and no participant received the same item twice.

3.3. Sentence task: materials

Following the design of the lexical tasks, material for the two language pairs consisted of a list of 90 sentences that could be either generally true (e.g., *Priests believe in God*; N = 45) or false (e.g., *Newborns drink beer*; N = 45). All items were simple sentences containing a transitive or ditransitive verb and its arguments. To ensure some variability between items, some items also contained negation, and some other a modal verb.

Crucially, in the SEN-F experiment, including French and LSF sentences, word order was always the same (congruent word order: SVO) in the spoken and in the signed version of each sentence and consistent with both grammars, whereas in the SEN-I experiment, with Italian and LIS sentences, word order was always different (incongruent word order: SVO/SOV) as required by each grammar. Nouns and verbs were selected so that no sign required an obligatory mouth component, and signs displaying multiple meanings or regional variations were avoided.

To select our items, we performed several pilots, starting with one with French hearing non-signers and one with Italian hearing non-signers. They were asked to judge the truth value of a wider list of sentences presented in written French or in written Italian, respectively, and, for each country, those sentences that received an unexpected rating by more than 7% of the participants were excluded. The reduced set of French sentences was adapted into LSF and was also evaluated by two Deaf LSF signers, and the reduced set of Italian sentences was adapted into LIS and evaluated by one Deaf LIS signer. Those sentences judged as not clear by our Deaf informants were also not included as final experimental items.

The two CODA actors who produced the lexical items were also filmed producing the 90 target sentences plus 4 additional training items, in the same three conditions as in the lexical task. For each language pair, the recording setting was the same as for the lexical items and the cutting procedure was identical.

For each language pair, the 90 target items were divided into three lists of 30 items each, 15 true and 15 false, so that each list had a similar number of transitive/ditransitive sentences, modals, and negations.

In their relevant language pair and for each condition, each participant received one different list, with lists counterbalanced across participants, so that all items appeared in all conditions and no participant received the same item twice.

3.4. Procedure

Both the LSF-French and the LIS-Italian experiments were implemented on Labvanced (Finger et al., 2017) and distributed online through a web link. The structure of the experiment was the same for the two language pairs. Questions and instructions, as well as informed consent forms, were all presented in written language (hence French or Italian, respectively).

Firstly, participants were asked a number of biographical questions. Only participants who declared to have at least one deaf signing parent were allowed to undertake the experiment.

Participants either took the lexical task followed by the sentence task, or the other way around, in a counterbalanced order. Each task was preceded by specific written instructions.

The instructions of the lexical task were to decide as fast and accurately as possible whether each noun produced in the video corresponded to an entity that one can "eat or drink". Participants responded with two keys: Q-M for French participants (AZERTY keyboard) and A-L for Italian participants (QWERTY keyboard). The location of the two keys is the same across the two types of keyboards. The match between the yes/no answer and the two answer keys varied across participants.

The instructions of the sentence task were to decide as fast and accurately as possible if the meaning of each sentence produced in the video was generally true or false. The match between the yes/no answer and the two keys was coherent with that of the lexical task.

As in the original experiment by Emmorey et al. (2012), within each task, participants received the three lists each in a separate block, i.e., there was a block per condition. The order of blocks was counterbalanced across participants, and each block began with 4 training items in the appropriate condition.

In both tasks, each trial began with a fixation point of 1000ms, followed by the video clip.

For both tasks, we recorded accuracy and reaction times (RTs). For items in the spoken language condition, RTs were assessed from the voice onset of the audio track (RTs-audio), calculated on the basis of the relative spectrogram visualized with PRAAT (Boersma & Weenink, 2022). For items in the sign language condition, RTs were calculated from the beginning of the video clip (RTs-video). For each item in the code-blending condition, two RTs were calculated, one from the voice onset of the audio track and one from the beginning of the video clip.

3.5. Data analysis

All statistical analyses were performed using R version 4.1.2 (R Core Team, 2021). We performed a separate analysis for each language pair and for each task. Before proceeding, we checked overall participants accuracy to remove participants with an error rate higher than 20%. No participant was excluded considering this criterion. Visual inspection of the French participants' data (Figures on OSF) clearly showed that one French participant was extremely slow in performing the task, being an outlier in almost all conditions. This participant was removed from the analysis.

Accuracy analysis was performed on all items[8], whereas RTs analysis considered only correct answers. For each task in each language pair, we performed two different RTs analyses. We considered RTs-audio to compare responses in the spoken language and in the code-blending conditions, whereas we considered RTs-video to compare responses in the sign language and in the code-blending condition.

RTs that were two standard deviations above or below the mean for each participant for each condition were discarded from the analysis. Considering the French experiments, 5.0% of RTs-video and 4.8% of RTs-audio were discarded in the lexical task, and 5.2% of RTs-video and 4.6% of RTs-audio in the sentence task. Considering the Italian experiment, this implied the elimination of 5.7% of RTs-video and 4.5% of RTs-audio in the lexical task and 4.7% of RTs-video and 4.9% of RTs-audio in the sentence task.

Accuracy data were analyzed with generalized mixed-effects regression (family: binomial, link function: logit; Jaeger, 2008) with condition as fixed factor (treatment coding with code-blending as the reference level, to compare accuracy in the sign language condition with accuracy in the code-blending condition and accuracy in the spoken language condition with accuracy in the code-blending condition).

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press

Table 1. Overall mean accuracy for semantic categorization decisions to LSF signs, French words, and LSF–French code-blends (LEX-F) and to LIS signs, Italian words, and LIS-Italian code-blends (LEX-I).

	Condition		
Experiment	Sign language M (SD)	Spoken language M (SD)	Code-blending M (SD)
LEX-F	90% (30%)	98% (15%)	98% (15%)
LEX-I	84% (36%)	98% (14%)	98% (15%)

RT data were analyzed with generalized mixed-effects regression (family: inverse gaussian, link function: identity, Lo & Andrews, 2015; see also Farhy et al., 2018). Condition was entered in the models as fixed factor, with code-blending coded as 0 in each analysis, and the other level (either sign language, for RTs-video or spoken language, for RTs-audio) as 1.

Both in the accuracy and in the RTs data analyses we used the maximal random effect structure justified for the design, i.e., 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.

Anonymized data, as well as the R scripts to perform the analysis are stored on the OSF page of the present paper (osf.io/z8smh/).

4. Results

4.1. Lexical task

Both French and Italian CODAs were highly accurate in the spoken language and in the code-blending conditions, and less accurate in the sign language condition (Table 1).

Specifically, for French CODAs, accuracy in the sign language condition (LSF) was significantly lower than accuracy in the code-blending condition (LSF + French) ($b = -1.59$, $z = -5.55$, $p < .0001$), whereas accuracy between the spoken language condition (French) and the code-blending condition did not significantly differ ($b = 0.05$, $z = 0.14$, $p = .8920$). As for Italian CODAs, accuracy in the sign language condition (LIS) was significantly lower than accuracy in the code-blending condition (LIS + Italian) ($b = -2.59$, $z = -2.66$, $p = .0078$), whereas accuracy in the spoken language condition (Italian) did not significantly differ with accuracy in the code-blending condition ($b = -0.16$, $z = -0.12$, $p = .9030$).

As for the RTs, for French CODAs RTs-video did not significantly differ between the sign language condition and the code-blending condition ($b = 12$, $t = 0.89$, $p = .37$), whereas RTs-audio were significantly slower in the spoken language condition compared to the code-blending condition (French + LSF) ($b = 161$, $t = 15.76$, $p < .0001$) (Figure 2).

Considering Italian CODAs, RTs-video were significantly slower in the sign language condition compared to the code-blending condition ($b = 240$, $t = 5.66$, $p < .0001$), and RTs-audio were significantly slower in the spoken language condition compared to the code-blending condition ($b = 94$, $t = 3.22$, $p = .0013$) (Figure 3).

4.2. Sentence task

Accuracy in the sentence task was higher in the spoken language and in the code-blending condition than in the sign language condition both for SEN-F and SEN-I, involving French and Italian CODAs, respectively (Table 2).

In particular, French CODAs' accuracy in the sign language condition (LSF) was significantly lower than accuracy in the code-blending condition (LSF + French) ($b = -2.11$, $z = -9.02$, $p < .0001$), whereas the difference in accuracy between the spoken language condition (French) and the code-blending condition was minimal and not significant ($b = -0.27$, $z = -1.00$, $p = .3160$). Likewise, Italian CODAs' accuracy in the sign language condition (LIS) was significantly lower than accuracy in the code-blending condition (LIS + Italian) ($b = -1.85$, $z = -6.07$, $p < .0001$), whereas accuracy in the spoken language condition (Italian) was higher than in the code-blending condition, but the difference was not significant ($b = 1.66$, $z = 1.82$, $p = .0691$).

Turning to RTs, in French CODAs RTs-video were significantly higher in the sign language condition (LSF) compared to the code-blending condition (LSF + French) ($b = 844$, $t = 9.51$, $p < .0001$), while RTs-audio did not differ between the spoken language condition (French) and the code-blending condition (French + LSF) ($b = 18$, $t = 0.57$, $p = .0570$) (Figure 4).

For Italian CODAs RTs-video were also significantly higher in the sign language condition (LIS) compared to the code-

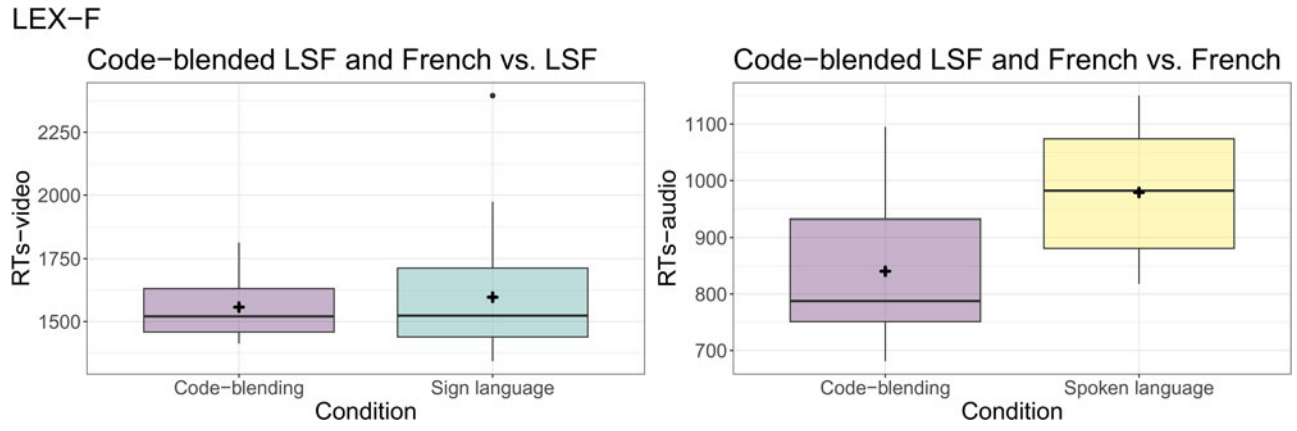


Figure 2. LEX-F: Box plots representing the distribution of by subject mean RTs-video (left) and RTs-audio (right) by condition (code-blending vs. sign language and code-blending vs. spoken language). The "+" symbol indicates the mean value by condition.

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press

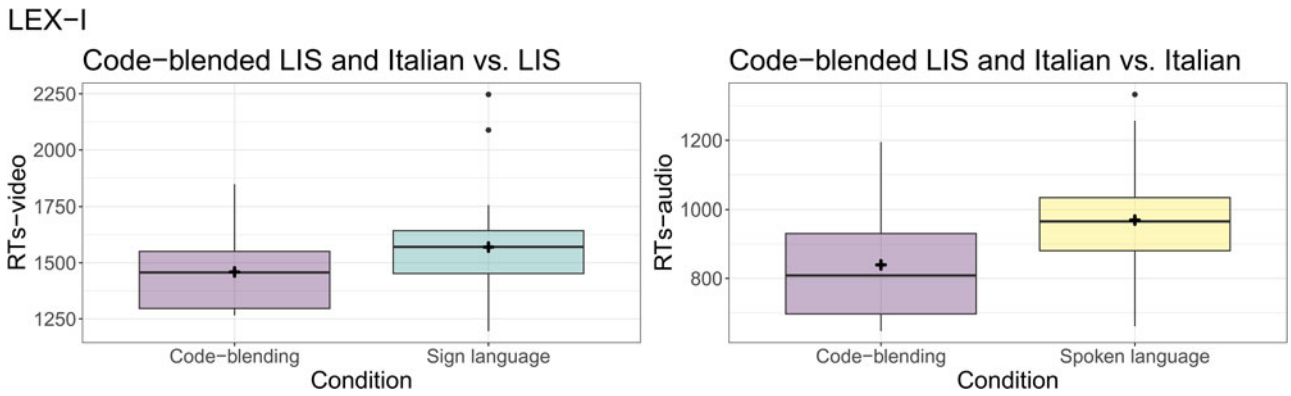


Figure 3. LEX-I: Box plots representing the distribution of by subject mean RTs-video (left) and RTs-audio (right) by condition (code-blending vs. sign language and code-blending vs. spoken language). The "+" symbol indicates the mean value by condition.

blending condition (LIS + Italian) (b = 704, t = 14.25, $p < .0001$), but RTs-audio were significantly lower in the spoken language condition (Italian) compared to the code-blending condition (Italian + LIS) (b = −157, t = −3.05, $p = .0023$) (Figure 5).

5. Discussion

The two lexical experiments that were performed with French and Italian CODAs substantially replicate the findings of Emmorey et al. (2012): in LEX-I there was always a code-blending advantage, both with respect to spoken language (Italian) and with respect to sign language (LIS), while in LEX-F the code-blending advantage was limited to code-blending vs. spoken language (French). Even if RTs for LSF–French code-blends were slightly lower than for LSF signs, there was no significant difference between the two conditions.

As for LEX-F, the advantage of code-blending over French is of primary importance, since French is the dominant language of the participants. In general, the dominance of the spoken language in CODAs is well known in the literature such that researchers argue that CODAs are a special type of heritage speakers (e.g., de Quadros, 2018; Polinsky, 2018). The spoken French dominance of our CODA participants is also confirmed more concretely by accuracy data, which show that participants were significantly less accurate in the sign language condition than in the spoken language or code-blending condition (Tables 1 and 2).

It is therefore reasonable to conclude that our data support the idea that double lexical access is not costly, but can rather be advantageous for bimodal bilinguals, and that processing a lexical item in two modalities beats the cost of double lexical access.

Table 2. Accuracy for truth value decisions to LSF sentences, French sentences, and LSF–French sentential code-blends (SEN-F) and to LIS sentences, Italian sentences, and LIS-Italian sentential code-blends (SEN-I).

	Condition		
Experiment	Sign language M (SD)	Spoken language M (SD)	Code-blending M (SD)
SEN-F	80% (40%)	95% (22%)	96% (19%)
SEN-I	82% (39%)	98% (15%)	96% (19%)

Turning to the sentence tasks, we found that blended sentences were evaluated faster than sign language sentences both in SEN-F and SEN-I. This can probably be explained by the fact that participants, being less confident in their sign language, relied more on speech when performing this more complex task. It is also possible that the lack of mouth movements in the signed videos, which were avoided on purpose in creating the signed stimuli given the obvious incompatibility with code-blending (see above, §3 including footnote 6) might have caused some extra difficulty in comprehension. This is because in spontaneous conversations most signed utterances are indeed supplemented with mouthings reproducing the equivalent words (or parts of them) in the spoken language (see e.g., Crasborn et al., 2008; Vinson et al., 2010; and § 3.2, note 7).

As for the crucial comparison between code-blending and spoken language the two experiments yielded divergent results.

Starting from SEN-I, we found that responses to code-blended items were significantly slower than responses to Italian-only items. This clearly suggests a cognitive cost in processing two non-congruent sentences on the fly.

In SEN-F, on the other hand, we found no significant difference between code-blending and spoken language (French). As is generally the case with null results, they are difficult to interpret. One possibility is that the experimental design was not accurate enough to detect an existing code-blend advantage with congruent sentences, a point to which we shortly come back.

With respect to our initial hypothesis, the combination of lexical alignment with either Lillo-Martin et al.'s or Branchini and Donati's proposals correctly predicts the results in the incongruent sentence experiment (SEN-I), while they diverge on the predictions for the congruent sentence experiment (SEN-F). In fact, the hypothesis including Lillo-Martin et al.'s proposal predicts a processing advantage for the blended condition over the unimodal (speech only) condition. This hypothesis is not corroborated by the experimental evidence. This in turn may receive an explanation under the hypothesis that includes Branchini and Donati's proposal. Under this second hypothesis, the lack of difference between the blend and the unimodal speech-only condition can be interpreted by assuming that the positive bias of lexical integration is balanced by the negative bias of processing two separate but similar syntactic structures.

A way to better understand the null result we found in SEN-F would be to run an analogous experiment involving a

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press



Figure 4. SEN-F: Box plots representing the distribution of by subject mean RTs-video (left) and RTs-audio (right) by condition (code-blending vs. sign language and code-blending vs. spoken language). The "+" symbol indicates the mean value by condition.

similar semantic categorization task with three signs/words utterances under the same three conditions (French, LSF, congruent code-blend), but this time involving simple lists of words, crucially not inserted into a syntactic structure. For example, we could provide lists such as *Berlin, Paris, Chicago* and ask participants to answer the question: *Is there an intruder?* As far as lexical access is concerned, these items would not differ from those of SEN-F (at least those without modals) in the code-blend condition. In this case, as in SEN-F, each lexical item would be synchronically paired twice in the code-blend condition. But this time, no syntactic structure needs to be processed in order to achieve the task. If we were to find an advantage of code-blend in this case, we would be in a better position to conclude that it is the processing of (congruent) syntactic structures that cancels the advantage associated with a double lexicalization.

A second direction for further research is to use more sensitive measures of processing mechanisms. It is a fact that experimental measures that exploit a single index of comprehension difficulty, like reaction times and accuracy, might not be reliable because slow responses may reflect slow processing, careful processing, or both (Xiang et al., 2014). Methodologies that allow to control for the tradeoff between processing speed and processing accuracy, as is the case with speed accuracy tradeoff methodology, might be the way to go.

Finally, the study of how code-blending is processed could not be complete without addressing the issue of how integrated blends are comprehended. It would be of great interest to assess with a similar experimental design whether integrated code-blends such as (11) present an advantage or not over their unimodal counterparts (12) and (13).

(11) It: Io *Code-blended (LIS-Italian) condition*
I
LIS: WIN
win
'I win.' (adapted from Branchini & Donati, 2016, p. 37)

(12) WIN *LIS condition*
'I won.'

(13) Vinco *Spoken Italian condition*
win-1SING
'I won.'

Remember that integrated blends like (11) are instances of incongruent blends (no alignment between redundant lexical items) where fragments belonging to the two languages are produced simultaneously and need to be integrated in a unique proposition/utterance. Since meaning is computed from structure, these are likely to involve one single syntactic structure, where items belonging to the two languages are integrated. In this

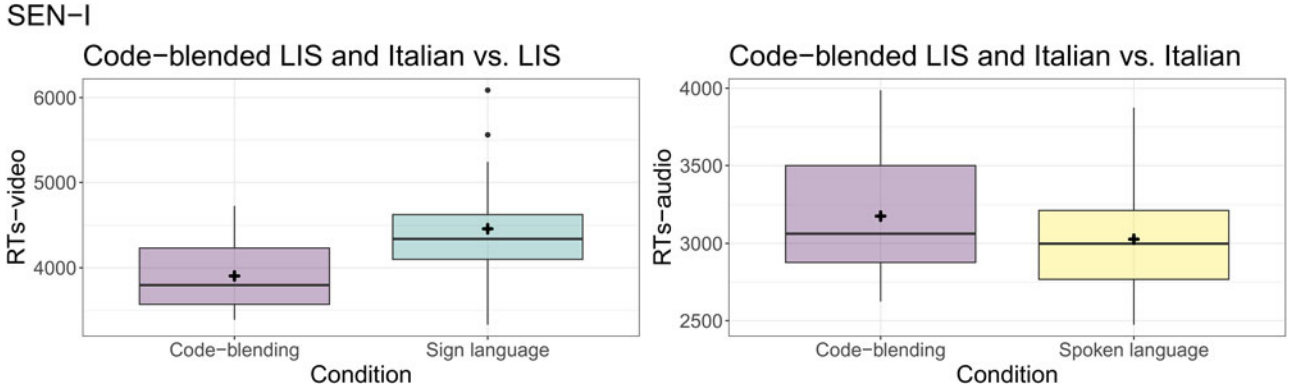


Figure 5. SEN-I: Box plots representing the distribution of by subject mean RTs-video (left) and RTs-audio (right) by condition (code-blending vs. sign language and code-blending vs. spoken language). The "+" symbol indicates the mean value by condition.

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press

case, even the proponents of the hypothesis that congruent and incongruent blends involve two syntactic structures (one corresponding to each simultaneous utterances) would posit a one and only structure, associated to the one and only proposition that is produced scattered along the two channels. If lexical misalignment is what makes incongruent code-blends difficult, we would expect these to be hard to process as well. If what makes sentential blends more difficult is the processing of two syntactic structures associated to two potentially independent simultaneous utterances/propositions, this difficulty should not hold here.

Be that as it may, we believe that studying the processing of blended utterances of various degrees of complexity, integration, and congruence might play a crucial role in answering long-standing questions about the interaction and accessibility of the two languages in bilingual individuals.

6. Conclusions

In this paper, we investigated the comprehension of bimodal bilingual utterances both at the lexical and the sentential level by conducting four on-line experiments. In the first two lexical semantic judgment experiments, we replicated a previous finding by Emmorey et al. (2012) with two new pairs of languages (LSF–French and LIS–Italian) – namely, that processing code-blended bimodal bilingual stimuli is not costly with respect to monolingual stimuli, but rather advantageous.

In the other two experiments, we investigated whether the same effect is also found at the sentential level. Results showed that in syntactically congruent language pairs (SEN-F, LSF–French) no difference was found between the code-blending condition and the spoken language condition, while an advantage was found for the code-blending condition over the sign language condition. In syntactically incongruent language pairs (SEN-I, LIS–Italian) a disadvantage was found for the code-blending condition over the spoken language condition, while an advantage was found over the sign language condition. These results call for further experimental studies of code-blending in sentence processing.

Data availability. The data that support the findings of this study are openly available in OSF (osf.io/z8smh/)

Acknowledgements. We thank our CODA consultants, Candy Prouhèze and Rita Sala. Moreover, we are grateful to Federica Loro and Alessandra Gentili for their help in data collection. We also thank all participants who took part in the study and those, Deaf and hearing, who took part in the various pilots we ran. We are grateful to the audience of the following conferences and seminars for their comments and suggestions: 35th Annual Conference on Human Sentence Processing (HSP 2022), Theoretical Issues in Sign Language Research (TISLR 14); Collège de France Seminar; Language Attitudes and Bi (dia)lectal Competence (LABiC 2022); Bilingualism Matters Research Symposium (BMRS 2022).

Authors' contributions. All authors conceived and designed the study; AJ and CB prepared the experimental materials; AJ implemented the tasks; BG analysed the data; all authors discussed the results; CD and BG drafted the manuscript; CG and CB revised the drafts. All authors have approved the manuscript before submission.

Notes

1. Notice that ASL is a null argument language (Koulidobrova, 2012), so the utterance in (1) is not to be considered incomplete in ASL. In general, all sign languages at least allow for some argument omission (see Jaber et al., 2022 for a typological account).

2 Examples of code-blending are transcribed using up to five lines. The spoken language data is given in the first line while the second line includes the corresponding glosses, when needed. Sign language data are given in the third line and the corresponding glosses are provided in the fourth line. On the first and third line, abbreviations of the language names appear at the beginning of the data line, e.g., "Can" for Cantonese, "HKSL" for Hong Kong Sign Language, « Engl » for English, « ASL » for American Sign language; « It » for Italian, and so on. Free translation of the entire utterance is provided on the last line. Following the conventions of sign linguistics research, the signing data are notated in small capitals. English is adopted to gloss signs with their closest possible meanings.

3 Of course, integrated semantically incongruent code blends are also syntactically incongruent, given that they display by definition no alignment of equivalent material.

4 An exception, in addition to LIS–Italian bilinguals discussed in the text, is HKSL (which is head final) and Cantonese (head initial), in the studies stemming from the Hong Kong research group on Deaf child and adult bilinguals: see Fung & Tang (2016) for a review.

5 A third logical possibility, suggested by an anonymous reviewer (whom we thank), is that two structures are generated only when needed: hence in incongruent blends (LIS–Italian), but not in congruent ones (French–LSF). This hypothesis would give rise to the same predictions as the one-structure/one or two linearizations hypothesis.

6 Considering the difficulties related to the recruitment of this special population, our goal was to reach and possibly outnumber the sample size of CODAs who took part in the lab-based lexical experiment of Emmorey et al. (2012) (N = 18, referred to as "early bimodal bilinguals") and who showed a code-blending facilitation effect.

7 Sign languages are produced not only through movements of the hands, but also through the use of many non-manual articulators, one of which is the mouth, through which lexical information can be encoded (see e.g., Pfau & Quer, 2010). Mouth movements are of two types, mouth gestures and mouthings. The former are not related to the spoken language and usually have an adjectival or adverbial function, the latter are derived from spoken language, and generally correspond to the mouthing of the first syllable(s) of the corresponding spoken word. Some signs are accompanied by an obligatory mouth gesture and we explicitly excluded these signs in our experiments. As for mouthings, their exact status within the phonological system of sign languages is controversial. Interestingly, some experimental evidence suggests that they do not belong to the core language system, but rather reflect a form of code-mixing (e.g., Giustolisi et al., 2017). If this is true, utterances containing mouthings would be instances of congruent code-blending. In our experiments, therefore, we explicitely asked the actors not to produce any mouthing in the signed condition, in order a) to be sure that the signed condition were genuinely unimodal, and b) to maximize the similarity with the blended condition, where of course mouthing was not possible given the simultaneous utterance of the spoken string. As one anonymous reviewer correctly points out, we might have lost some naturality in the sign and blended conditions with this manipulation, but we believe that there was no choice in order to obtain a controlled experiment.

8 In the French lexical task (LEX-F), there was an implementation error that caused the loss of some items for some participants (details in the script on OSF). For this reason, we removed 52 responses corresponding to 2% of the data.

References

Baker, A. E., & Van den Bogaerde, B. (2008). Code-mixing in signs and words in input to and output from children. In Plaza-Pust, G. and Morales-López, E. (eds.) *Sign Bilingualism: Language development, interaction, and maintenance in sign language contact situations* (pp. 1–27). London: John Benjamins.

Bishop, M. (2011). HAPPEN CAN'T HEAR: An Analysis of Code-Blends in Hearing, Native Signers of American Sign Language. *Sign Language Studies 11*, 205–240.

Bishop, M., & Hicks, S. (2005). Orange Eyes: Bimodal Bilingualism in Hearing Adults from Deaf Families. *Sign Language Studies 5*, 188–230.

Bishop, M., Hicks, S., Bertone, A., & Sala, R. (2007). Capitalizing on simultaneity: Features of bimodal bilingualism in hearing Italian native signers. *Multilingualism and Sign Languages: From The Great Plains to Australia*. 79–118.

Blanco-Elorrieta, E., Emmorey, K., & Pylkkänen, L. (2018). Language switching decomposed through MEG and evidence from bimodal bilinguals.

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press

Proceedings of the National Academy of Sciences 115(39), 9708–9713. doi: 10.1073/pnas.1809779115
Boersma, P., & Weenink, D. (2022). Praat: doing phonetics by computer Version 6.3.01, retrieved 26 November 2022 from http://www.praat.org/.
Branchini, C., & Donati, C. (2016). Assessing lexicalism through bimodal eyes. *Glossa: a journal of general linguistics 1*(1), 48. doi: 10.5334/gjgl.29
Branchini, C., & Mantovan, L. (2020). *A Grammar of Italian Sign Language (LIS)*. Edizioni Ca' Foscari – Venice University Press. http://doi.org/10.30687/978-88-6969-474-5
Chomsky, N. (1970). Remarks on nominalization, in Jacobs, R., & Rosenbaum, P. (eds.), *Readings in English transformational grammar* (pp. 184–221). Waltham (MA): Ginn.
Crasborn, O., van der Kooij, E. D., Woll, B., & Mesch, J. (2008). Frequency distribution and spreading behavior of different types of mouth actions in three sign languages. *Sign Language and Linguistics 11*, 45–67. doi: 10.1075/sll.11.1.04cra
Crepaldi, D., Amenta, S., Pawel, M., Keuleers, E., & Brysbaert, M. (2015). SUBTLEX-IT. Subtitle-based word frequency estimates for Italian. *Annual Meeting of the Italian Association for Experimental Psychology*, Rovereto, 10–12 September 2015.
de Quadros, R. M. (2018). Bimodal Bilingual Heritage Signers: A Balancing Act of Languages and Modalities. *Sign Language Studies 18*(3), 55–384.
de Quadros, R. M., Davidson, K., Lillo-Martin, D., & Emmorey, K. (2020). Code-blending with depicting signs. *Linguistic aproaches to bilingualism 10*(2), 290–308.
Dijkstra, T., & Kroll, J. F. (2005). Bilingual visual word recognition and lexical access. *Handbook of bilingualism: Psycholinguistic approaches*, *178*, 201.
Donati, C., & Branchini, C. (2013). Challenging linearization: Simultaneous mixing in the production of bimodal bilinguals. In Roberts, I., & Biberauer, T. (eds.), *Challenges to linearization*, (pp. 93–128). Berlin: Mouton De Gruyter.
Emmorey, K., Borinstein, H., Thompson, R., & Gollan, T. (2008). Bimodal bilingualism. *Bilingualism: Language and Cognition*, *11*(1), 43–61. doi: 10.1017/S1366728907003203
Emmorey, K., Petrich, J., & Gollan, T. (2012). Bilingual processing of ASL–English code- blends: The consequences of accessing two lexical representations simultaneously. *Journal of Memory and Language*, *67*(1), 199–210. doi: 10.1016/j.jml.2012.04.005
Farhy, Y., Veríssimo, J., & Clahsen, H. (2018). Do late bilinguals access pure morphology during word recognition? A masked-priming study on Hebrew as a second language. *Bilingualism: Language and Cognition*, *21*(5), 945–951. doi: 10.1017/S1366728918000032
Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: A Unified JavaScript Framework for Online Studies. In *International Conference on Computational Social Science IC2S2*, 2016–2018.
Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of memory and language*, *51*(1), 1–22. doi: 10.1016/j.jml.2004.01.004
Fung, C. H. M., & Tang, G. (2016). Code-blending of functional heads in Hong Kong Sign Language and Cantonese: A case study. *Bilingualism: Language and Cognition*, *19*(4), 754–781. doi: 10.1017/S1366728915000747
Giezen, M. R., Blumenfeld, H. K., Shook, A., Marian, V., & Emmorey, K. (2015). Parallel language activation and inhibitory control in bimodal bilinguals. *Cognition 141*, 9–25. doi: 10.1016/j.cognition.2015.04.009
Giustolisi, B., Mereghetti, E., & Cecchetto, C. (2017). Phonological blending or code mixing? Why mouthing is not a core component of sign language grammar. *Natural Language and Linguistic Theory 35*, 347–365. doi: 10.1007/s11049-016-9353-9
Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale, & S. J. Keyser (Eds.), *The view from building 20* (pp. 111–176). The MIT Press.
Ikeda, S. (1998). The Process of Translation Between English Words and Japanese Words, *The Journal of General Psychology*, *125*(2), 117–128. doi: 10.1080/00221309809595536
Jaber, A., Donati, C., & Geraci, C. (2022). On the properties of null subjects in sign languages: the case of French Sign Language (LSF). *The Linguistic Review 39*(4), 655–686. doi: 10.1515/tlr-2022-2100
Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446. doi: 10.1016/j.jml.2007.11.007
Kanto, L., Huttunen, K., & Laakso, M. L. (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), 242–260. doi: 10.1093/deafed/ens071
Koulidobrova, E. (2012). Parallelism Revisited: The Nature of the Null Argument in ASL as Compared to the Romance-Style Pro. *Sign Language and Linguistics 15*(2), 259–270. doi: 10.1075/sll.15.2.07kou
Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and cognition*, *9*(2), 119–135. doi: 10.1017/S1366728906002483
Lillo-Martin, D., de Quadros, R. M., & Chen Pichler, D. (2016). The development of bimodal bilingualism: Implications for linguistic theory. *Linguistic Approaches to Bilingualism 6*(6), 719–755. doi: 10.1075/lab.6.6.01lil
Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology 6*, 1171. doi: 10.3389/fpsyg.2015.01171
Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics 41*(5), 791–824. doi: 10.1515/ling.2003.026
Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Similarity within and across languages. *Journal of Psycholinguistic Research 37*(3), 141–170. doi: 10.1017/S1366728903001068
Morford, J., Wilkinson, E., Villwock, A., Piñar, P., & Kroll, J. (2011). When deaf signers read English: Do written words activate their sign translations? *Cognition 118*, 286–292. doi: 10.1016/j.cognition.2010.11.006
Muysken, P. (2000). *Bilingual speech: a typology of code-mixing*. Cambridge: Cambridge University Press.
New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics 28*(4), 661–677. doi: 10.1017/S014271640707035X
Ormel, E., Hermans, D., Knoors, H., & Verhoeven, L. (2012). Cross- language effects in written word recognition: The case of bilingual deaf children. *Bilingualism: Language and Cognition 15*(2), 288–303. doi: 10.1017/S1366728911000319
Petitto, L. A., Katerlos, M., Levy, B., Gauna, K., Tetreault, K., & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: Implications for the mechanisms underlying early bilingual language acquisition. *Journal of Child Language 28*, 453–496. doi: 10.1017/s0305000901004718
Pfau, R., & Quer, J. (2010). Nonmanuals: Their grammatical and prosodic roles. In Brentari, D (ed.), *Sign Languages* (Cambridge Language Surveys, pp. 381–402). Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511712203.018
Polinsky, M. (2018). Sign Languages in the Context of Heritage Language: A New Direction in Language Research. *Sign Language Studies*. *18*, 412–428. doi: 10.1353/sls.2018.0009
R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
Rinaldi, P., & Caselli, C. (2014). Language development in a bimodal bilingual child with cochlear implant: A longitudinal study. *Bilingualism: Language and Cognition 17*(4), 798–809. doi: 10.1017/S1366728913000849
Rinaldi, P., Lucioli, T., Sanalitro, C., & Caselli, M. C. (2021). Code-blending in bimodal bilingual deaf children. In *Rivista di Psicolinguistica Applicata*, vol. *21-1*, 47–60. doi: 10.19272/202107702004
Shook, A., & Marian, V. (2012). Bimodal bilinguals co-activate both languages during spoken comprehension. *Cognition*, *124*(3), 314–324. doi: 10.1016/j.cognition.2012.05.014
Vinson, D. P., Thompson, R. L., Skinner, R., Fox, N., & Vigliocco, G. (2010). The Hands and Mouth Do Not Always Slip Together in British Sign Language: Dissociating Articulatory Channels in the Lexicon. *Psychological Science*, *21*(8), 1158–1167. https://doi.org/10.1177/0956797610377340
Xiang, M., Dillon, B., Wagers, M., Liu, F., & Guo, T. (2014). Processing covert dependencies: an SAT study on Mandarin wh-in-situ questions. *Journal of East Asian Linguistics 23*, 207–232. doi: 10.1007/s10831-013-9115-1
Zwitzerlood, I. (2012). Classifiers. In Pfau, R., Steinbach, M., & Woll, B. (eds.). *Sign Language: An International Handbook* (pp. 158–186). Berlin, Boston: De Gruyter Mouton.

https://doi.org/10.1017/S1366728923000901 Published online by Cambridge University Press