Intra-sentential Code Switching at the Syntax-prosody Interface*

Mandy Cartner
Julia Horvath
Tel Aviv University
0009-0008-2743-6089; mandycartner@mail.tau.ac.il
0000-0002-5648-4847; horvath@tauex.tau.ac.il



© the authors

Reception: September 14, 2025
Acceptance: May 14, 2026
Publication: October 6, 2026

Recommended citation: Cartner, Mandy & Horvath, Julia. 2026. Intra-sentential Code Switching at the Syntax-prosody Interface. *Catalan Journal of Linguistics* 25(2): 63-91. <https://doi.org/10.5565/rev/catjl.562>

Abstract

Grammatical patterns of intra-sentential code-switching (ICS) provide a testing ground for architectural questions in generative linguistics. ICS shows that bilinguals are able to derive sentences containing material in more than one language. We investigate the nature of a minimal unit for ICS, which must remain unilingual, and propose to derive it from already established properties of grammar, specifically the mapping from syntax to phonology. We report data from three acceptability studies on Hebrew-English bilinguals, investigating ICS patterns at the CP edge, the DP edge, and with pronominal subjects. Our results point to a prosodically-defined minimal unit for ICS, which does not adhere to the boundaries established by Phase theory, contra López et al.'s (2017) theory of "code-switching by Phase".

Keywords: code switching; phases; prosodic structure; bilingualism; English; Hebrew

Resum. *Canvi de codi intraoracional a la interfície sintaxi-prosòdia*

Els patrons gramaticals del canvi de codi intraoracional (ICS) permeten contrastar qüestions sobre l'arquitectura del llenguatge en lingüística generativa. L'ICS mostra que els bilingües poden generar oracions que contenen material en més d'una llengua. Investiguem la naturalesa d'una unitat mínima per a l'ICS, que ha de romandre monolingüe, i proposem derivar-la de propietats gramaticals prèviament establertes, concretament el mapatge de la sintaxi a la fonologia. Presentem

* We thank Gilly Hoffman Even and Nathan Lewin for recording the materials used in our experiments, and the audiences at CGG 33, NELS 56, Tel Aviv University, and UMass Amherst for their feedback and discussion. This research was supported by a Collaborative Research Grant from the Cukier-Goldstein-Goren Center for Mind and Language (MILA), awarded to PIs Julia Horvath and Aya Meltzer-Asscher (2021–2023).

dades de tres estudis d'acceptabilitat amb bilingües d'hebreu i anglès, i investiguem patrons d'ICS al marge del CP, al marge del DP i amb subjectes pronominals. Els nostres resultats apunten cap a una unitat mínima definida prosòdicament per a l'ICS, que no s'ajusta als límits establerts per la teoria de les fases, en contra de la teoria de «canvi de codi per fases» de López et al. (2017).

Paraules clau: canvi de codi; fases; estructura prosòdica; bilingüisme; anglès; hebreu

Table of Contents

1. Introduction
2. Experiments 1A-B: The CP edge
3. Experiments 2A-B: The DP edge
4. Experiments 3A-B: Pronominal subjects
5. General discussion
6. Conclusions
References

1. Introduction

Intra-sentential code switching (ICS) is the use of more than one language within the boundaries of a single sentence, exemplified in (1a). ICS is common among fluent bilinguals. Alongside the various social factors that influence the practice of ICS, studies find that ICS is grammatically constrained (e.g., Poplack 1980; Di Sciullo, Muysken & Singh 1986; Mahootian 1993; Myers-Scotton 1993; MacSwan 1999, 2006; Muysken 2000; González-Vilbazo 2005; González-Vilbazo & López 2011; López, Alexiadou & Veenstra 2017; López 2020). For example, while the sentence in (1a) with a switch between a Spanish auxiliary and an English progressive is considered acceptable, Spanish-English bilinguals are reported to reject (1b), with a switch between a Spanish auxiliary and an English perfective.

(1) a. Los ciudadanos están *supporting the program.* Spanish/*English*
The citizens are

b. *Tú habías *told that story*.
You had

(López, Alexiadou & Veenstra 2017)

Deriving such grammaticality patterns is the focus of theoretical work on ICS. While early research on ICS patterns often made use of constraints specific to ICS, we pursue the null hypothesis that bilinguals' patterns of ICS derive from general properties of bilinguals' grammar (Mahootian 1993; MacSwan 1999). This approach, in turn, suggests that ICS patterns can provide valuable insight into the organization of bilingual grammar, and grammar in general.

MacSwan (1999) develops a model of ICS which is embedded in the Minimalist Program (MP) of generative grammar (Chomsky 1995). MacSwan posits that bilinguals are equipped with two traditional lexicons to which parametric differences are attributed, while the syntactic component lacks language-specific rules and

thus can be shared across a bilingual's two languages. Lexical items may be drawn from either lexicon and inserted into a single derivation, introducing features which must be checked for convergence, no different from the derivation of a unilingual sentence.

While MacSwan (1999) motivates a single syntactic component, he argues that bilinguals must represent distinct PFs, one for each language. He argues that the PF component cannot be unified because PF rules are ordered/ranked with respect to one another, and these orders are language-specific (Bromberger & Halle 1989). In order for each ordering to be maintained, it must be encapsulated in a language-specific PF. Relatedly, research into the phonetics and phonology of bilinguals motivates the assumption that bilinguals have two phonological systems, as they successfully maintain two sets of phonetic targets and phonological processes, even in ICS (Caramazza et al. 1973; Olson 2019; Gosselin 2022).

Under this architectural assumption, the derivation of an ICS sentence must be divided between two PF components. A crucial question to explore then is: Can any unit of derivation be divided as such? Or are there units that must be sent off to a single PF component? This would determine a minimal unit for CS, a unit that cannot involve CS internal to it but must be in a single language.

We pursue this question within the framework of Distributed Morphology (DM) (Halle & Marantz 1993). Under DM, the information associated with lexical items of previous models is distributed across various components of grammar, such that the objects inserted into syntax are abstract and lack phonological exponence, making them potentially non-language-specific. Thus, the choice between one language or another in bilingual grammar may only affect the derivation at spell-out, where an abstract syntactic representation is assigned language-specific exponence. While ICS patterns are likely influenced by parametric differences across a bilingual's two languages, commonly represented as contrasting formal-feature specifications within the MP, this study focuses on whether and how properties of the syntax-phonology interface and the mapping of an abstract syntactic representation to language-specific PFs affect ICS patterns.

López (2020) develops a DM model for bilingual grammar to account for ICS patterns, but postulates a single shared PF. He argues for this based on cases where PF rules associated with one language apply to, or affect the processing of, material in a bilingual's other language. Examples cited include a Spanish impoverishment rule for clitic sequences that bilinguals may sometimes apply to Catalan; Welsh-English bilinguals exhibiting processing difficulty when a Welsh mutation rule is applied to an English word in the wrong environment (Vaughan-Evans et al. 2014); light verbs imposing the prosodic structure and constituent order of their language on a complement in the other language (González-Vilbazo & López 2012).

However, we argue that the existence of such cases may not provide sufficient support for the single-PF assumption; it could also be the result of grammatical transfer due to language contact, or of co-activation across separate grammatical systems. For example, the application of a Spanish impoverishment rule in Catalan could be the result of transfer across separate PF systems, i.e., adoption of the Spanish rule into the Catalan PF system of the speaker, rather than neces-

sarily evidence for a shared inventory of PF rules. Thus, the examples in López (2020) do not necessarily motivate the stronger claim that bilinguals represent a single PF, which would face empirical challenges. Beyond the facts noted above supporting MacSwan (1999)'s model, a single PF component in bilingual grammar predicts competition between vocabulary items across languages. For example, due to the Elsewhere Principle (Halle & Marantz 1993), in Hebrew-English grammar, the Hebrew exponent of 'with' should always win against the English exponent when this preposition appears with a pronominal complement (contrary to the fact), since Hebrew specifies a contextually restricted form with pronouns, while the English form functions as the elsewhere exponent. Accordingly, we assume that bilinguals represent language-specific PF systems, as this assumption avoids the empirical challenges faced by a single-PF model.

Given these architectural assumptions, we investigate whether *prosodic structure*, established at the syntax-phonology interface, influences ICS patterns. Our investigation focuses on function words, which are assumed to be prosodically dependent (Selkirk 1996; Elfner 2015), in configurations where syntactic and prosodic structure are expected to align. In the next section, we review previous work by MacSwan (1999) and López et al. (2017), each suggesting a PF-based restriction on ICS, and present our alternative hypothesis. In section 1.2. we discuss some important methodological considerations. Section 2, 3 and 4 present the design and results of three studies we conducted. Section 5 discusses our conclusions.

1.1. Mapping ICS at PF

Based on his assumption that bilinguals' represent two distinct PFs, MacSwan (1999) proposes that ICS within a single head is impossible, as each head must be interpreted by a language-specific PF component. In particular, he argues that since phonological rules are sensitive to inflectional affixes (Halle & Mohanan 1985), a complex syntactic head must be sent to one language-specific PF, e.g., PF of L_x. If this head hosts material in L_y, this L_y material is unidentified by the L_x PF system, and cannot be parsed. Thus, a stem and an inflection affix must be in the same language, and no ICS can occur between them. This accounts for Poplack's (1980)1980 observation that a switch between a root and bound morpheme is banned, as in (2).

(2) *Juan está *eat*-iendo Spanish/*English*
Juan be eat-DUR
'Juan is eating.'

Moreover, MacSwan utilizes this proposal to explain several acceptability patterns in Spanish-Nahuatl ICS. For example, he reports that switching between an auxiliary and an inflected verb is unacceptable, and argues that this is the result of inflected verbs raising to T^0 for feature-checking, resulting in a complex syntactic head. Similarly, switching between a Spanish determiner and a Nahuatl noun is also unacceptable, which MacSwan argues is the result of Spanish determiners

triggering noun raising for agreement purposes, resulting in a complex D^0. Within his model, these complex syntactic heads must be sent to a single language-specific PF component for interpretation. Since any material in the other language is uninterpretable at the other language's PF, code-switching is impossible.

More recently, López, Alexiadou & Veenstra (2017) proposed that ICS is only possible between *phasal spell-out domains* and prohibited within one. Phasal spell-out domains constitute cycles of derivation that are transferred to the interfaces before further structure is built (Chomsky 2001). According to the prevalent view, Phases are CP and vP (and sometimes DP), and their complements (TP and VP, respectively) undergo transfer. The transferred material, i.e., the Phase's spell-out domain, becomes inaccessible to further syntactic probing, while the Phase head and specifier (the Phase *edge*) remain active in the next higher Phase (Chomsky 2001). Phasal and spell-out boundaries are illustrated for the sentence *Maria studied laws* in Figure 1.

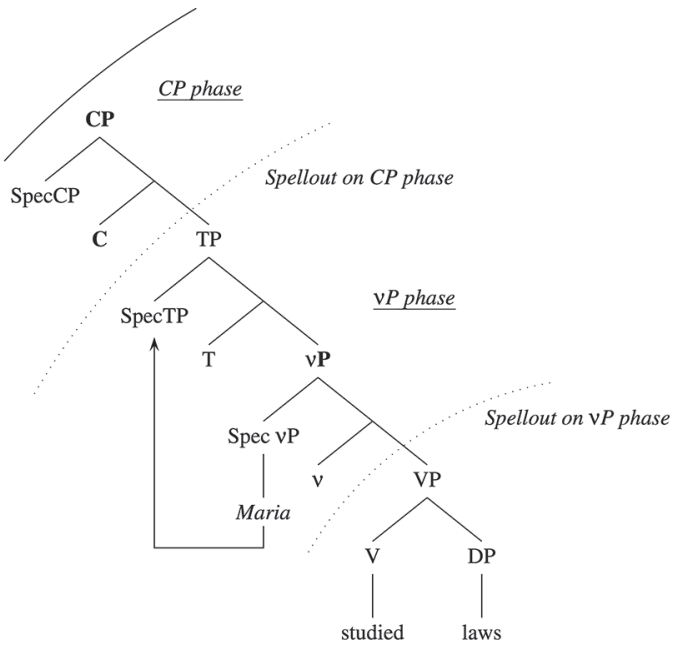


Figure 1. An illustration of Phases and spell-out for the derivation of *Maria studied laws*. (Kratzer & Selkirk 2007: 104).

López et al. (2017) propose that since the whole spell-out domain undergoes transfer to the interfaces as one unit, it is sent off to a single, language-specific phonological system and hence must be realized all in one language. They predict that material within a spell-out domain must match in language, such that ICS is possible only across spell-out boundaries. Thus, they account for the contrast in (1), repeated in (3). In (3b), the auxiliary and the verb are spelled-out together, as part of the CP complement domain, and cannot involve code-switching. On the other hand, the authors suggest that the progressive in (3a) introduces a Phase boundary between T and the *-ing* verb, based on the locative PP analysis for progressives (Bolinger 1971), making code-switching between the auxiliary and progressive possible.

(3) a. Los ciudadanos están *supporting the program.* Spanish/*English*
The citizens are

b. *Tú habías *told that story*.
You had

(López, Alexiadou & Veenstra 2017)

However, as noted by the authors, patterns of ICS at the CP Phase challenge the proposal that ICS is constrained by spell-out domains. They report that switching between a *wh*-phrase and a TP is acceptable (4a), while switching between a complementizer and a TP is not (4b).

(4) a. *Ich weis nicht*, [$_{CP}$ *welches Buch* [$_{TP}$ Juan compró]] *German*/Spanish
I know not which book Juan bought.
'I don't know which book Juan bought.'

(González-Vilbazo & López 2013)

b. *Eduardo *denkt* [$_{CP}$ que [$_{TP}$ *Elena schreibt sich im*
Eduardo thinks that Elena Registers SELF in.DEF.DAT
sekretariat]]
secretary's office
'Eduardo thinks that Elena registers at the secretary's office.'

(González-Vilbazo 2005)

To account for the contrast in (4), they suggest that transfer is sensitive to the feature content of the elements at the Phase edge (based on Ott 2011). In the embedded interrogative (4a), C is marked with an interpretable [+Q] feature which must be checked at the higher Phase, ensuring that C stays behind and is spelled-out separately from its complement TP, allowing for ICS. However, in declarative clauses (4b), C allegedly bears no uninterpretable or clause-typing features that require further checking, and thus it is transferred along with its complement, blocking ICS between the CP edge and complement.

Note however that their proposed distinction may be confounded: The *wh*-phrase in (4a) is lexically restricted, while the complementizer in (4b) is a bare function word. In this study, we explore the possibility that the source of the con-

trast in (4) is prosodic, and specifically, that ICS between prosodically dependent elements is restricted.

Prosodic structure refers to the abstract hierarchical organization of linguistic material into phonological constituents (e.g., prosodic words and phrases), which serve as the domain for rhythmic and phonological computations. This structure is widely assumed to emerge from the interaction of syntactic and phonetic constraints (Nespor & Vogel 1986). In particular, a large body of work posits systematic correspondences between syntactic and prosodic constituency (Selkirk 1986, 2009, 2011; Truckenbrodt 1995, 1999; Wagner 2010; Elfner 2015; Bennett & Elfner 2019; Büring & Truckenbrodt 2021). The assumption of syntax-prosody correspondence is based on the tendency of phrasal stress and contextually restricted phonological processes to respect syntactic boundaries. For example, gemination in Tuscan Italian only applies when two words belong to the same syntactic phrase (e.g., VP-adverb and verb), and does not apply across syntactic phrase boundaries (e.g., sentential adverb and verb) (Truckenbrodt 1995).

Crucially, theories of prosodic parsing suggest that only lexical elements trigger the formation of an independent prosodic phrase (also called 'phonological phrase'), while functional elements are parsed along with some lexical material (e.g., Selkirk 1996; Elfner 2015), unless they are prosodically prominent. Thus, the complementizer in (4b), but not the *wh*-phrase in (4a), is expected to be prosodically dependent. We investigate the possibility that the prosodically dependent nature of function words restricts ICS options. Accordingly, we propose the following hypothesis, to be tested in the present study:

(5) *ICS by prosodic phrase*
ICS cannot occur within a prosodic phrase; it is possible only at prosodic phrase boundaries.

Broadly, the *ICS by prosodic phrase* hypothesis predicts a systematic asymmetry between lexical and functional material in ICS: a lexical phrase may be freely switched, whereas functional elements must match the language of the lexical material with which they form a prosodic phrase (on the functional-lexical distinction in ICS see also Belazi, Rubin & Toribio 1994; Grimstad, Lohndal & Åfarli 2014; Joshi 1982; Myers-Scotton 1993). Under standard theories of syntax-prosody correspondence described above, the prosodic attachment of a functional element is expected to align with its syntactic constituency. In the absence of additional factors, this yields the prediction that a functional head should match the language of its complement (e.g., D with N).

Our proposal that prosodic structure constrains ICS, rather than syntactic structure directly, is motivated by the assumption that the PF component is the locus of language-specific information, while the syntactic component is abstract and shared across languages within a bilingual's grammar (MacSwan 1999; Gosselin 2022). In the current study, we test the *ICS by prosodic phrase* hypothesis in configurations in which syntactic and prosodic structure are expected to align; particularly, at the CP edge, the DP edge, and with pronominal subjects. However, various phonetic and

information structure factors can influence prosodic structure, potentially resulting in mismatches between syntactic and prosodic constituency (Selkirk 1996). Such cases could provide an important further test for our proposal as they would allow us to tease apart the contribution of abstract structure from the phonetic properties of an utterance. We set aside these cases for future work.[1]

We report data from 3 acceptability studies on Hebrew-English bilinguals, comparing the predictions of MacSwan (1999)'s theory, which we label *ICS by head*, our proposed theory *ICS by prosodic phrase*, and López et al. (2017)'s theory *ICS by Phase*. Experiments 1A-B target ICS at an embedded CP, manipulating the prosodic size and feature-content of the CP edge element (Section 2), experiments 2A-B target ICS at the DP edge (Section 3) and experiments 3A-B target ICS with pronominal subjects, manipulating the Phase status of the surrounding clause. Before presenting the studies, Section 1.2 motivates the methodology used in our experiments.

1.2. Methodology

Early work on ICS patterns was based on informal acceptability judgements, which often resulted in empirical disagreements between researchers. For example, some researchers claim that ICS between a complementizer and TP is unacceptable (Belazi, Rubin & Toribio 1994; López, Alexiadou & Veenstra 2017), and others claim that it is (Joshi 1982; Di Sciullo Muysken & Singh 1986; MacSwan 2000). While judgements on ICS vary based on external factors, such as speakers' language background, fluency, and presentation modality (González-Vilbazo et al. 2013; Badiola & Sande 2018; Koronkiewicz & Ebert 2018), these factors are often not disclosed in non-experimental work on ICS, leaving discrepancies across studies unaccounted for. González-Vilbazo et al. (2013) provide an exploration of different methodologies for ICS research, stressing the importance of the experimental design, participant selection, and the explicit reporting of methodology. In accord with the conclusions of this work and others (e.g., Gullberg, Indefrey & Muysken 2009), we investigate grammatical patterns of ICS experimentally.

This research utilizes acceptability judgment tasks using auditory, rather than written stimuli. This mode of presentation avoids the orthographic problem arising in Hebrew-English ICS, as each language uses an opposite writing direction. Auditory stimuli were recorded by a balanced Hebrew-English bilingual, who was naive to the purpose of the experiment, and who was instructed to use neutral intonation, maintain a consistent speaking rate, and avoid pauses or focus prosody at the site of code-switching. Each sentence was recorded twice, and the second

1. One motivation for our focus on syntactically determined abstract prosodic structure rather than phonetic realization is our concern that ICS itself may give rise to pauses or accentuation that would not otherwise be expected. Accordingly, additional linking assumptions would be required to relate abstract prosodic structure to specific phonetic cues in ICS, in order for these to be properly manipulated in an experimental setting.

production was used in the experiment, as it is typically more fluent and stable. All stimuli were noise-reduced and normalized using Audacity.

Lastly, in order to understand participants' language background, all participants completed the Bilingual Language Profile (BLP) questionnaire (Birdsong, Gertken & Amengual, 2012) which consists of questions about one's language history, use, proficiency, and attitudes. Each participant's responses are scored for the calculation of their relative language dominance score, ranging from -218 for most Hebrew dominant, to +218 for most English dominant.

2. Experiments 1A-B: The CP edge

Experiments 1A and 1B are auditory judgment tasks, designed to test whether Hebrew-English bilinguals' ICS patterns at the CP edge are modulated by prosodic structure, and/or complement spell-out. The two experiments differ only in the direction of the switch: all items in Experiment 1A switched from Hebrew to English, and all items in Experiment 1B switched from English to Hebrew. Switch direction was held constant within each experiment for design parsimony and variance control; our theoretical hypotheses do not predict systematic differences between switch directions.

These experiments are designed to examine López et al.'s (2017) claim that code-switching at the CP edge is modulated by the feature-content of C, such that *wh*-phrases may mismatch the language of the embedded clause, while declarative complementizers cannot, as a function of spell-out (discussed above in relation to (4a,b)). This prediction is compared against an alternative hypothesis, *ICS by prosodic phrase*, which predicts ICS at the CP edge to be modulated by prosodic size: ICS with small, prosodically dependent function words is predicted to be constrained, while lexical phrases forming independent prosodic units may be switched freely. These hypotheses critically contrast in their predictions for bare *wh*-words: While *ICS by Phase* predicts bare *wh*-words to align with lexically restricted *wh*-phrases, due to their [+Q] feature, *ICS by prosodic phrase* predicts them to align with complementizers, due to their prosodic status as bare function words. Finally, MacSwan's (1999) *ICS by head* predicts all patterns of ICS tested in Experiments 1A-B to be acceptable, since material at the CP edge should not incorporate any surrounding material into a single head.

2.1. Methods

Participants. Participants were recruited via social media. 101 participants completed the task, until data from 96 (48 in each task) suitable Hebrew-English bilinguals were collected. 5 participants were removed because they reported that they were also fluent speakers of Russian. Table 1 summarizes Experiments 1A-B's participants' assigned dominance score and their age of acquisition (AoA) for each language, which is their self-reported age of first regular exposure. On average, participants were more Hebrew dominant, with the large majority residing in Israel at the time of testing. The average age of participants was 26.5 (*Std. 5.6*). We also

asked participants about their code-switching habits: 83.3% of the participants in 1A and 77% in 1B self-report that they switch between languages in the course of a conversation "sometimes" or "often", while the rest answered that they "rarely" or "never" do so.

Table 1. BLP data for Experiments 1A-B. The Dominance scale ranges from -218 to 218, and the self-reported proficiency scales range from 1 to 6

	Experiment 1A				Experiment 1B			
	Avg.	**Std.**	**Min.**	**Max.**	**Avg.**	**Std.**	**Min.**	**Max.**
Dominance	-48.32	45.52	-140.01	82.12	-64.33	39.31	-133.12	25.94
Hebrew AoA	0.73	1.74	0	6	0.71	2.60	0	6
English AoA	3.54	3.01	0	10	4.63	3.53	0	12
Hebrew proficiency	5.76/6	0.73	2	6	5.85/6	0.45	4	6
English proficiency	5.11/6	0.88	3	6	5.10/6	1.00	2	6

Materials. The materials were pre-recorded by a bilingual speaker of Hebrew and English, as described in Section 1.2. All experimental items included a switch at the edge of an embedded clause. In each experiment, 24 sets included embedded *wh*-interrogatives, in which two factors were manipulated: whether the *wh*-phrase matched the language of its clause (*C+TP*: the language of C and TP *match* or *mismatch*) and what is its type (*wh*-type: *bare wh*-word or a *lexical wh*-phrase). The extraction site was counterbalanced across sets, such that half included subject extraction and half direct object extraction. In 12 other sets (in each experiment), the language of a declarative complementizer was manipulated (*C+TP*: *match* or *mismatch*). Example items can be seen in Table 2. Experimental items were accompanied by an additional 36 filler items, all involving ICS.

Table 2. Example item sets from Experiment 1A, switching from Hebrew to English

Embedded declaratives (n=12)		
	C+TP mismatch	*ha-melṣarit* *hodi'a* *še-* the restaurant is closing soon the-waitress announced that
	C+TP match	*ha-melṣarit* *hodi'a* that the restaurant is closing soon the-waitress announced
Embedded interrogatives (n=24)		
bare	C+TP mismatch	*ha-mazkira* *badka* *mi* is supposed to come in today the-secretary checked who
	C+TP match	*ha-mazkira* *badka* who is supposed to come in today the-secretary checked
lexical	C+TP mismatch	*ha-mazkira* *badka* *eize* *mitlamed* is supposed to come in today the-secretary checked which intern
	C+TP match	*ha-mazkira* *badka* which intern is supposed to come in today the-secretary checked



Figure 2. Acceptability scale used in the experiments.

Procedure. The experiment was hosted on the online platform PCIbex farm (Zehr & Schwarz 2018), and completed by each participant on their personal device. The instructions informed participants that they will hear sentences that mix Hebrew and English and will be requested to provide their impression of how natural each sentence sounds, on a 6pt scale. The experiment began with 5 practice items. At each trial, the recorded sentence started playing automatically and was followed by a rating scale, as in Figure 2. Participants made a choice on the scale using the number keys on their keyboard, from 1 to 6. There was no time limit for rating, and participants could replay the audio file. Each trial was followed by a 1000ms break. After completing the judgment task, participants filled out the BLP questionnaire.

2.2. Analysis and Results

Embedded declaratives. Participants' ratings for the embedded declaratives item sets were analysed using Bayesian cumulative probit mixed-effects models (using the brms package in R, Bürkner 2017), one for each experiment (1A and 1B).[2] The models included *C+TP* as a predictor, and random intercepts and slopes for participants and items. Weakly informative priors were used throughout. The model was estimated with 6 chains, each running for 5,000 iterations (1,000 warmup). We consider parameters whose 95% credible interval (CrI) does not contain 0 to indicate reliable effects – that is, when the data provide evidence for a positive or negative effect rather than one that could plausibly be absent.

For both experiments (1A & 1B), the models provided evidence for negative *C+TP* effects, such that sentences in which the complementizer mismatched the embedded clause were rated poorly compared to sentences in which they matched. The point estimate for this effect is larger in Experiment 1A. Mean ratings by condition are presented in Figure 3. Posterior values for the estimates and their errors, along with 95% CrI's, are provided in Table 3.

2. The results and analysis files, along with the full set of materials, are available in the following OSF repository: <https://osf.io/3xkwu/overview?view_only=bc23eb027de34191bf755ad2f9baafc2>.

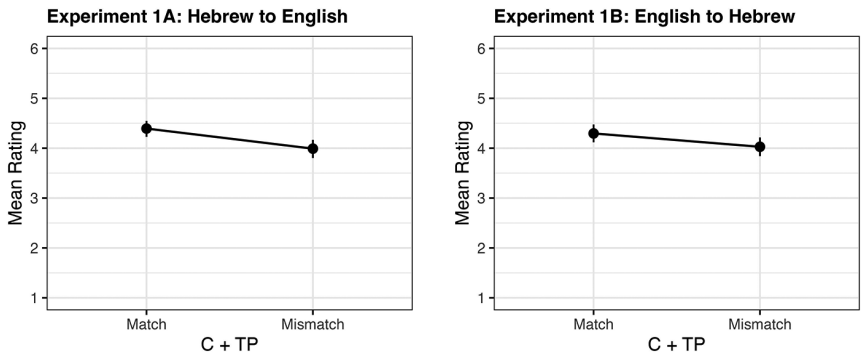


Figure 3. Mean acceptability ratings for *embedded declaratives* in Experiments 1A-B. Error bars show 95% bootstrapped confidence intervals for the mean rating.

Table 3. Bayesian cumulative probit mixed-effects models fit to the *embedded declaratives* data from Experiments 1A-B

	Estimate	Est. Error	95% CrI lower	95% CrI upper
Experiment 1A				
C+TP	-0.19	0.10	-0.38	-0.00
Experiment 1B				
C+TP	-0.11	0.06	-0.22	-0.00

Embedded interrogatives. Participants' ratings for the *embedded interrogative* item sets were analysed using Bayesian cumulative probit mixed-effects models, one for each experiment. The models included *C+TP* and *Wh Type* as predictors. All other model details were identical to the models used for the *embedded declaratives* dataset.

For both experiments (1A & 1B), the models provided evidence for *C+TP* effects, such that participants preferred sentences in which the *wh*-element matched rather than mismatched the language of the embedded clause. Crucially, this effect interacted with *Wh Type*, such that the preference for the *wh*-element and the embedded TP to match was only present with bare *wh*-words, and not with lexically restricted *wh*-phrases. The models also provided evidence for a main effect of *Wh Type*, indicating higher acceptability ratings with lexically restricted *wh*-phrases compared to bare *wh*-words, which likely stems from the sharp decrease in acceptability for bare *wh*-words in the *C+TP mismatch* condition. This effect, as well as the interaction, seem to be more robust in Experiment 1A than 1B. Mean ratings by condition are presented in Figure 4. Posterior values for the estimates and their errors, along with 95% CrI's, are provided in Table 4.

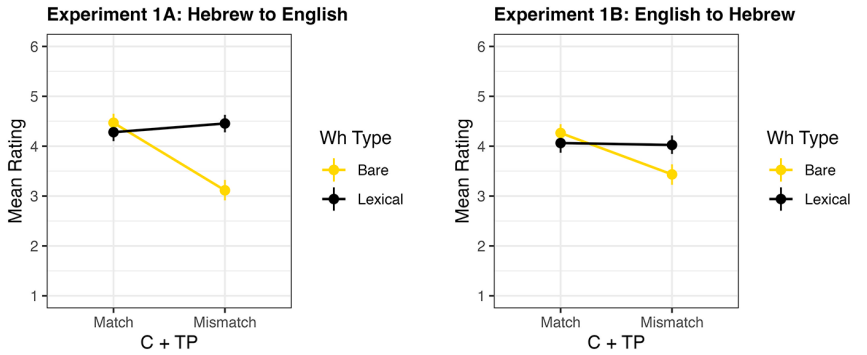


Figure 4. Mean acceptability ratings for *embedded interrogatives* in Experiments 1A-B. Error bars show 95% bootstrapped confidence intervals for the mean rating.

Table 4. Bayesian cumulative probit mixed-effects models fit to the *embedded interrogative* data from Experiments 1A-B

	Estimate	Est. Error	95% CrI lower	95% CrI upper
Experiment 1A: Hebrew to English				
Wh Type	0.23	0.05	0.14	0.32
C + TP	-0.24	0.05	-0.34	-0.13
Wh Type x C + TP	0.33	0.05	0.22	0.43
Experiment 1B: English to Hebrew				
Wh Type	0.08	0.05	-0.01	0.17
C + TP	-0.16	0.05	-0.26	-0.05
Wh Type x C + TP	0.16	0.05	0.06	0.26

2.3. Discussion

The results of Experiments 1A-B demonstrate that Hebrew-English bilinguals' preferences for ICS within an embedded CP vary by the prosodic size of the CP edge element: With complementizers and bare *wh*-words, we found a preference for the CP edge to match the language of the TP, rather than mismatch it. Contrarily, with lexically restricted *wh*-phrases, this preference was nullified and both ICS between the CP edge and complement, as well as ICS before the CP edge, were similarly acceptable.

This pattern of acceptability supports the proposed theory of ICS by prosodic phrase, as it is selective of size: Only bare functional material at the CP edge must match the language of the embedded material, its apparent host, while lexically restricted phrases are independent and free to switch (see Büring & Truckenbrodt 2021 on the prosodic phrasing of bare vs. lexical *wh*-phrases). This pattern can be captured neither by the *ICS by head* (MacSwan 1999) nor by the *ICS by Phase*

(López et al. 2017) theories. First, MacSwan's proposal does not predict these ICS patterns to be constrained, because elements at the CP edge do not form a single head with any surrounding material. In fact, MacSwan casts doubt on early reports that ICS between a complementizer and TP is unacceptable (Belazi, Rubin & Toribio 1994), and does not attempt to accommodate them within his theory.

López et al. (2017)'s theory *ICS by Phase* does not capture the reported pattern of acceptability, either. López et al. propose that the spell-out of complementizers and *wh*-phrases differs based on the feature-content of C, capturing the contrast they observe between complementizers and *wh*-phrases. Accordingly, *bare wh*-words would be expected to pattern with *lexically restricted wh*-phrases, since both specify a [+Q] feature, ensuring they are spelled-out with the next higher Phase, and should thus match the language of the matrix verb and may mismatch the embedded TP. However, we observe the opposite pattern: Bilinguals prefer bare *wh*-words to match the language of TP, similarly to complementizers.

More broadly, the patterns of acceptability reported in Experiments 1A-B do not support the boundaries predicted by the standard MP Phase theory. We observe that bare function words at the CP edge must match the language of the embedded TP, despite the assumption that the CP edge and complement are spelled-out separately. This is also surprising under the *ICS by prosodic phrase* hypothesis: given the assumption that spell-out feeds prosodic structure, prosodification across a spell-out boundary would be blocked. However, prosodification of C with its complement is expected under theories of syntax-prosody correspondence, and supported by cases of complementizer allomorphy that is conditioned by properties of the embedded subject (Bošković 2016). This is discussed further in the General Discussion (Section 5).

3. Experiments 2A-B: The DP edge

Experiments 2A-B are auditory judgment tasks, designed to test the acceptability of switching between a determiner and a NP. As in 1A-B, the two experiments differ only in the direction of the switch (2A: Hebrew→English; 2B: English→Hebrew).

Assuming that DP is a Phase (Chomsky 2000; Bošković 2005), *ICS by Phase* (López et al. 2017) predicts switching between a determiner and its complement NP to be acceptable, as they are assumed to be spelled-out separately. In fact, *ICS by Phase* predicts determiners to obligatorily match the language of the higher Phase with which they are spelled-out. On the other hand, under the *ICS by prosodic phrase* hypothesis, determiners are expected to match the language of the element with which they are prosodically parsed, given their status as prosodically dependent function words (Nespor & Vogel 1986; Elfner 2015). In line with standard syntax-prosody correspondence theories, prosodic constituency normally reflects syntactic constituency, so determiners are parsed together with their NP complement. This assumption is supported, for example, by the prevalence of phonologically conditioned allomorphy between determiners and their complement, such as the case of the English indefinite article which is dependent on the initial segment at the NP edge. We thus predict a preference for determiners to match the

language of the NP, contrasting with *ICS by Phase*, which predicts determiners to match the language of the higher Phase.

Finally, MacSwan (1999)'s *ICS by head* does not predict determiners in Hebrew-English to be constrained. MacSwan argues that ICS between a determiner and NP is constrained insofar as the determiner agrees with the noun, which triggers noun raising resulting in a complex D. Since neither the English nor Hebrew determiners tested in Experiments 2A-B agree with N, *ICS by head* predicts these to be freely switched.

3.1. Methods

Participants. Participants were recruited via social media. 96 Hebrew-English bilinguals (48 in each task) completed the task. See Table 5 for a summary of data collected in the BLP questionnaire. On average, participants were more Hebrew dominant, with the large majority residing in Israel. The average age of participants was 27.52 (*Std. 5.08*). Moreover, 71% of the participants in 2A and 79% in 2B self-report that they switch between languages during the course of a conversation "sometimes" or "often", while the rest answered that they "rarely" or "never" code-switch.

Table 5. BLP data for Experiments 2A-B. The Dominance scale ranges from -218 to 218, and the self-reported proficiency scales range from 1 to 6

	Experiment 2A				Experiment 2B			
	Avg.	**Std.**	**Min.**	**Max.**	**Avg.**	**Std.**	**Min.**	**Max.**
Dominance	-53.26	32.58	-119.51	36.59	-57.07	31.07	-124.7	25.15
Hebrew AoA	0.64	0.97	0	5	0.62	1.68	0	9
English AoA	2.28	2.66	0	7	3.66	2.87	0	9
Hebrew proficiency	5.92/6	0.26	4	6	5.64/6	0.91	3	6
English proficiency	5.21/6	0.8	3	6	5.07/6	0.91	3	6

Materials. The materials were pre-recorded by a bilingual speaker of Hebrew and English, as described in Section 1.2. In 24 sets, in each experiment, two factors were manipulated: whether a determiner matched the language of a following noun phrase (*D+NP match* or *mismatch*) and whether that noun phrase matched the language of the following verb (*NP+V match* or *mismatch*). The critical NPs were all multi-word (with PP modifiers), to avoid participants construing them as borrowing, and items began with a language-mismatched adjunct to prevent the determiner from being the sole mismatched element. The critical DP was always the clause's subject, not an internal argument, because internal arguments differ across Hebrew and English in ways that could influence acceptability ratings (e.g., Hebrew has accusative DOM). Focusing on subjects also avoids potential syntax-prosody mismatches in the parsing of determiners, which have been observed with internal arguments (Silber-Varod, 2014). 24 items from another task (Experiments

3A-B) were used as filler items for this experiment, in addition 12 other filler items, all involving ICS.

Table 6. Example item sets from Experiment 2B, switching from English to Hebrew

Determiner-NP (n=24)		
NP+V match	D+NP match	In my opinion, *ha-nose šel ha-xibur dey šanui be-maxloket* the-topic of the-essay bit lies in-debate
	D+NP mismatch	In my opinion, the *nose šel ha-xibur dey šanui be-maxloket* topic of the-essay bit lies in-debate
NP+V mismatch	D+NP match	In my opinion, *ha-nose šel ha-xibur* is a little controversial the-topic of the-essay
	D+NP mismatch	In my opinion, the *nose šel ha-xibur* is a little controversial topic of the-essay

Procedure. The details of the procedure were identical to Experiments 1A-B.

3.2. Analysis and Results

Participants' ratings were analysed using Bayesian cumulative probit mixed-effects models, one for each experiment. The models included *D+NP* and *NP+V* as predictors. All other model details were identical to the models used in Experiments 1A-B.

In both experiments (2A & 2B), the models did not detect reliable effects of *NP+V*, indicating similar acceptability ratings when the noun (subject) and verb match in language and when they mismatch. Crucially, the models provided evidence for *D+NP* effects, such that participants preferred sentences in which the determiner matched rather than mismatched the language of the noun phrase. In both experiments, this effect interacted with *NP+V*, such that the preference for the determiner and the NP to match was stronger when the NP also matched the following verb, compared to when they mismatched. This interaction was more robust in Experiment 2A than in Experiment 2B. The *D+NP* effect was also smaller overall in Experiment 2B. Mean ratings by condition are presented in Figure 5. Posterior values for the estimates and their errors, along with 95% CrI's, are provided in Table 7.

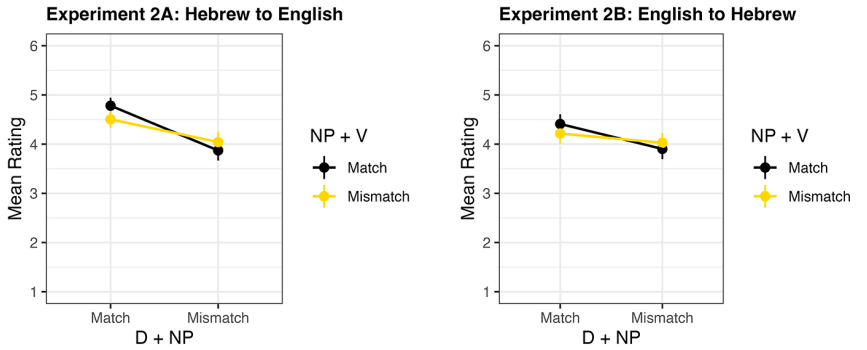


Figure 5. Mean acceptability ratings in Experiments 2A-B. Error bars show 95% bootstrapped confidence intervals for the mean rating.

Table 7. Bayesian cumulative probit mixed-effects models fit to the data from Experiments 2A-B

	Estimate	Est. Error	95% CrI lower	95% CrI upper
Experiment 2A: Hebrew to English				
NP + V	-0.04	0.05	-0.14	0.07
D + NP	-0.28	0.07	-0.41	-0.14
NP + V x D + NP	0.11	0.04	0.04	0.18
Experiment 2B: English to Hebrew				
NP + V	-0.03	0.05	-0.13	0.08
D + NP	-0.13	0.05	-0.23	-0.03
NP + V x D + NP	0.07	0.04	0.00	0.15

3.3. Discussion

The results of Experiments 2A-B demonstrate that Hebrew-English bilinguals prefer a determiner to match, rather than mismatch, the language of its complement NP. This pattern is consistent with the *ICS by prosodic phrase* hypothesis: like other bare function words such as complementizers and bare *wh*-words (Experiment 1), determiners cannot be freely switched. Importantly, we observe that the element controlling the determiner's language is its complement NP, consistent with the predictions of syntax-prosody correspondence theories, under which prosodically dependent functional heads are grouped with their lexical complements. In our stimuli, this direction of prosodic alignment is likely reinforced by the fact that the material preceding the determiner does not belong to the core clause. Future work could investigate whether switching preferences for determiners differ in configurations where syntactic and prosodic constituency may diverge.

On the other hand, the preference for determiners to match their complement NP is unpredicted by *ICS by Phase*: Since the determiner and embedded NP are spelled-out separately, ICS between them should be acceptable. Moreover, *ICS by head* does not account for this pattern, as the determiner constitutes a separate head, which should be free to switch.

In addition, we find some evidence suggesting that the preference for a determiner to match the language of its complement NP is stronger when this NP matches rather than mismatches the language of the TP. This interaction could potentially reflect an additional preference for D and TP to match, which would be consistent with *ICS by Phase*, assuming that D is spelled-out with TP. However, since participants display a robust preference for D to match NP, implementing this interaction within Phase theory would require D to be spelled-out both with its complement and with the higher TP. Because of this, and since the interaction term was rather weak, it likely reflects a general decrease in sensitivity in sentences that include multiple points of code-switching, which occurs in the *NP+V mismatch* conditions but not the *NP+V match* conditions. Moreover, as pointed out by an anonymous reviewer, the *D+NP match, NP+V match* condition does not include CS internal to the core clause, which might inflate its ratings relative to the other conditions. This potential confound could create the larger *D+NP* contrast observed in the *NP+V match* conditions.

Moreover, the preference for a determiner to match the language of the NP seems stronger in Experiment 2A (Hebrew→English) compared to 2B (English→Hebrew). If this difference is reliable, it might reflect a contrast between English and Hebrew determiners: While both are small function words, Hebrew determiners are often analyzed as clitics (Danon, 1996; Wintner, 1998), potentially strengthening the preference for them to match the host NP.[3] Note, however, that each switch direction was tested with a separate participant group, confounding a direct comparison between directions.

Lastly, it is important to note that previous work claims that ICS between a determiner and a noun is acceptable (Di Sciullo, Muysken & Singh 1986; Parafita Couto & Gullberg 2019). While this might reflect variation between different language-pairs, previous studies usually report ICS patterns for single word NPs. Our design intentionally targeted ICS between a determiner and a multi-word NP, to avoid ICS being interpreted as borrowing. This should be explored in an experiment that directly compares single and multi-word ICS within the DP.

4. Experiments 3A-B: Pronominal subjects

In order to investigate further the relationship between cyclic spell-out and a prosodic restriction on ICS, we ran Experiments 3A-B – auditory judgement tasks – focused on ICS patterns with pronominal subjects. ICS between a pronominal sub-

3. Alternatively, given that effect sizes are also larger in Experiment 1A than 1B, this could reflect a general trend of increased sensitivity in ICS from Hebrew to English, compared to the opposite direction.

ject and a verb is generally considered degraded (Timm 1975; Gumperz 1977; van Gelderen & MacSwan 2008). In particular, Koronkiewicz (2014) ran acceptability judgement tasks on Spanish-English bilinguals, comparing code-switching after weak or strong subject pronouns, the latter being modified, coordinated, stressed, or in a cleft/hanging topic construction. He reports that the ban on switching between a pronominal subject and a verb is more robust when the pronoun is weak compared to when it is strong. This is consistent with *ICS by prosodic phrase*, suggesting that this constraint may reflect the pronoun's status as a bare function word, prosodically dependent on the verb.

In Experiments 3A-B we further investigate the ban on ICS between a pronominal subject and verb, by comparing code-switching patterns in embedded clauses that are either finite or non-finite, such that only the former constitutes an independent Phase. This allows us to test whether the pattern of *ICS by prosodic phrase* emerging in the current study is consistent with Phase theory, despite our previous findings that are incompatible with phasal complement spell-out. In finite clauses (CP Phases), Phase theory predicts a spell-out boundary between the matrix verb and the embedded subject. If prosodification respects spell-out boundaries, that subject should be divorced from the main clause and must prosodify along with the embedded verb, which will block ICS between them. On the other hand, in the non-finite clauses we test, there is no spell-out boundary between the matrix verb and the embedded subject, which should allow prosodification of a pronominal subject to the matrix verb, and thus enable code-switching between that pronoun and the embedded verb. In other words, a contrast between the two types of clauses in ICS options is predicted if we assume that cyclic spell-out feeds prosodification, such that prosodic phrases cannot cross spell-out domains.

4.1. Methods

Participants. Participants were recruited via social media. 96 Hebrew-English bilinguals (48 in each task) completed the study. The participants are the same as the ones in Experiments 2A-B. See section 3.1, and in particular Table 5, for information about the participants.

Materials. The materials were pre-recorded by the same bilingual speaker of Hebrew and English as Experiments 2A-B, in the same session. The experimental items all included an embedded clause with a pronominal subject. In 24 sets, in each experiment, two factors were manipulated: (a) whether the embedded clause is finite or a non-finite small clause with an accusative marked subject (*Tense*: *finite* or *nonfinite*) and; (b) whether the pronominal subject matched the rest of the embedded clause in language (*pron+V*: *match* or *mismatch*).

Table 8. Example item sets from Experiment 3B, switching from English to Hebrew. Brackets represent Phase boundaries

Determiner-NP (n=24)		
finite	pron+V mismatch	The teacher saw [that they *meramim ba-bxina*] cheat in.the-exam
	pron+V match	The teacher saw [*še-hem meramim ba-bxina*] that-they cheat in.the-exam
nonfinite	pron+V mismatch	The teacher saw them *meramim ba-bxina* cheat in.the-exam
	pron+V match	The teacher saw *otam meramim ba-bxina* them cheat in.the-exam

The critical pronoun was always plural, to avoid homonymy across languages which might create confusion regarding the switch location (Hebrew nominative 3sg.M sounds like 'who' and Hebrew nominative 3sg.F sounds like 'he'). The verb in the finite clause was always in present tense to keep it morphologically identical to the non-finite verb. Material was constructed to be natural with both tenses, but there is still an unavoidable semantic contrast that accompanies the tense manipulation. 24 items from another task (Experiments 2A-B) were used as filler items for this experiment, in addition to 12 other filler items, all involving ICS.

Procedure. The details of the procedure were identical to Experiments 1A-B.

4.2. Analysis and Results

Participants' ratings were analysed using Bayesian cumulative probit mixed-effects models, one for each experiment. The models included *pron+VP* and *Tense* as predictors. All other model details were identical to the models used in Experiments 1A-B.

In both experiments (3A and 3B), the models revealed robust interaction terms between *pron+V* and *Tense*, reflecting a crossing pattern. In finite clauses, participants preferred the embedded pronominal subject to match the *embedded* verb in language, rather than mismatch it and match the matrix verb. In the non-finite clauses, the opposite pattern is observed: participants prefer the embedded pronominal subject to match the *matrix* verb, and mismatch the embedded verb in language. In Experiment 3A, no reliable main effects were detected for either *Tense* or *pron+V*, reflecting the crossing interaction. In Experiment 3B, the model detected a reliable effect of *Tense*, such that the *non-finite* condition received higher ratings on average than the *finite* condition. However, this effect was qualified by the interaction, indicating that it does not generalize across the levels of the *pron+V* condition. Mean ratings by condition are presented in Figure 6. Posterior values for the estimates and their errors, along with 95% CrI's, are provided in Table 9.

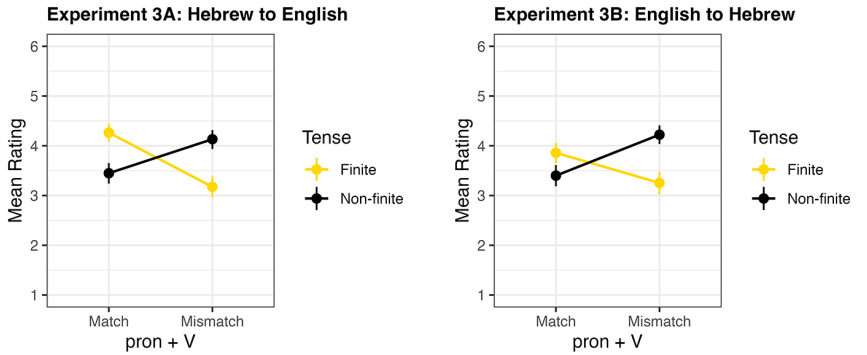


Figure 6. Mean acceptability ratings in Experiments 3A-B. Error bars show 95% bootstrapped confidence intervals for the mean rating.

Table 9. Bayesian cumulative probit mixed-effects models fit to the data from Experiments 3A-B

	Estimate	Est. Error	95% CrI lower	95% CrI upper
Experiment 3A: Hebrew to English				
Tense	0.04	0.06	-0.08	0.16
pron + V	-0.08	0.05	-0.19	0.02
Tense x pron + V	0.36	0.06	0.24	0.48
Experiment 3B: English to Hebrew				
Tense	0.10	0.04	0.03	0.19
pron + V	0.04	0.05	-0.06	0.13
Tense x pron + V	0.28	0.05	0.17	0.38

4.3. Discussion

The results of Experiments 3A-B demonstrate that Hebrew-English bilinguals' preferences for ICS with a pronominal subject are modulated by the Phase status of the surrounding clause: In finite embedded clauses, participants prefer a pronominal subject to match the language of the embedded verb, rather than mismatch it and appear in the language of the matrix clause. On the other hand, the opposite pattern is observed in the non-finite, small clauses, where participants preferred a pronominal subject to match the matrix rather than the embedded clause-mate verb.

These results further support our hypothesis that ICS is prosodically restricted, such that prosodically dependent functional material must match the language of its host. This pattern stands in contrast to the findings of Experiments 2A-B, where lexical subjects were free to mismatch the language of the following verb. Taken together, acceptability patterns across Experiments 2A-B and 3A-B highlight the

relevance of the lexical-functional distinction in ICS: pronominal subjects are restricted, while lexical subjects may be switched. Furthermore, the contrasting pattern across clause types observed in Experiments 3A-B is compatible with the idea that phase-by-phase spell-out feeds prosodification, such that a functional element cannot depend on material across a phasal spell-out boundary.

A remaining question is why the subject-pronouns of the non-finite clauses must match the language of the matrix, rather than the embedded verb. The lack of a CP boundary above the pronoun should not influence its ability to prosodify alongside the embedded material, predicting optionality between the two ICS patterns. It is possible that this preference reflects a particular syntactic input underlying prosodic structure, forcing the pronoun to be prosodified to the left, as the small clause subject might need to raise to the matrix VP for exceptional accusative Case. This predicts small-clause subjects to pattern with direct objects, which we are currently testing.

5. General discussion

We report Hebrew-English bilinguals' acceptability judgments for ICS sentences, collected across 3 auditory judgement studies testing possible ICS junctures, each comprising two experiments: One targeting ICS from Hebrew to English, and the other ICS from English to Hebrew. We argue that spell-out domains inadequately predict the set of possible ICS junctures, contrary to the proposal raised in López et al. (2017). Moreover, our results cannot be accounted for under MacSwan's (1999) proposal that ICS is impossible within a syntactic head. Instead, we have shown that ICS at the CP edge, DP edge, and with pronominal subjects is preferred when it takes place between separate prosodic phrases, while prosodically dependent functional material should match the language of its host.

In addition, we suggest that the patterns reported in MacSwan (1999) as well as those of López et al. (2017) could be accommodated under the present *ICS by prosodic phrase* proposal. First, since a head is always contained within a single prosodic phrase, MacSwan's (1999) proposal predicts a proper subset of the patterns our *ICS by prosodic phrase* predicts. MacSwan motivates restrictions on ICS within complex syntactic heads involving various function words, such as determiners, pronouns, conjunctions, auxiliaries and prepositions – which are all expected under *ICS by prosodic phrase*, as well. Moreover, the ICS patterns which motivate López et al. 's (2017) *ICS by Phase* all involve bare function words, such as auxiliaries or bound morphemes. The finding that these elements are constrained to match the language of material within their spell-out domain, we suggest, would also be accounted for under *ICS by prosodic phrase*, given the assumption that prosodification applies at each cycle of spell-out. For example, López et al. (2017) note that CS following a progressive auxiliary is acceptable, unlike CS following a perfective auxiliary (exemplified in (3)) because the former, but not the latter, introduces a Phase boundary between T and V. The observed distinct ICS behaviour of the two auxiliaries could receive a plausible account also under the present proposal: If prosodification applies to phasal spell-out domains, the progressive

auxiliary would be parsed separately from the verb, allowing CS between them, whereas the perfective auxiliary forms a single prosodic phrase with the verb, blocking a switch.

We derive the constraint our study motivated, *ICS by prosodic phrase*, from the assumption that bilingual grammar represents separate PF components. Due to the fact that the boundaries of prosodic phrases affect language-specific phonological and morphological rules, we hypothesize that the material within each prosodic phrase must be drawn from a single, language-specific, Vocabulary component. "Vocabulary" refers to the inventory of phonological exponence assigned to abstract syntactic terminals at the interface of the syntactic derivation with PF, within the architecture of DM.

A crucial question is how can this constraint, *ICS by prosodic phrase*, be enforced within the general architecture of DM applied to bilingual grammar, without utilizing ICS-specific stipulations. This depends on independent assumptions regarding the order of operations at PF, which are under debate (Pak 2008; Embick 2010; Arregi & Nevins 2012; Schuhmann & Putnam 2021). In particular, if we assume that Vocabulary Insertion (VI) takes place before prosodification, the ban on ICS within a prosodic phrase may be implemented as a filter: A code-switched string prosodified into a single phrase will not be able to be parsed by the phonology at later PF stages, resulting in a crash, much like MacSwan (1999) proposed for syntactic heads.

On the other hand, under the assumption that the assignment of prosodic structure precedes VI, *ICS by prosodic phrase* may be executed as a restriction on Vocabulary access: Once a structure is prosodified, only one language-specific Vocabulary component may be accessed at each prosodic phrase, blocking the possibility of inserting exponents from different Vocabularies in a single phrase, i.e., blocking ICS. Prosodic structure may precede VI within a theory of prosodification that depends only on syntactic bracketing, and not on exponence, such as Selkirk (1996). The broader debate about the relative ordering of PF operations remains unsettled, but bilingual data on ICS offer a unique testing ground, as they may shed light on the kind of information PF operations are sensitive to, and in what respects they are language-specific.

In particular, much research has explored cases where prosodic structure diverges from syntactic constituency, due to factors such as language-specific parametric variation, phonological size or weight, focus stress, or other information-structure effects on pitch accent (Büring 2013; Samek-Lodovici 2005; Wagner & Watson 2010). In the present study, we focused on relatively uncontroversial configurations and neutral intonation and information structure in which syntax and prosody are expected to align. Investigating ICS patterns in contexts where prosodic and syntactic structure conflict would provide insight into whether ICS is sensitive to finer-grained prosodic distinctions and their phonetic correlates (cf. Koronkiewicz 2014), and accordingly, which level(s) of representation it is sensitive to. Such work could clarify the relationship between abstract prosodic structure – potentially established before VI – and realization, which further depends on phonological properties of vocabulary material as well as semantic and discourse properties.

Notably, the prosodic restrictions observed in our study do not align with the boundaries defined by *complement spell-out*, a prevalent assumption of Phase theory, since we found a preference for prosodically dependent function words to match the language of material *across* a complement spell-out boundary, both for CPs and DPs. This is surprising under the assumption that prosodification applies to spell-out domains, blocking the possibility of a prosodic phrase across a spell-out boundary. Instead, these ICS patterns are compatible with *full Phase spell-out*, wherein the assignment of prosodic structure is sensitive to the boundaries of full Phases, rather than phasal complements, allowing material at the Phase edge to be prosodified along with the Phase complement.

These findings favouring full Phase spell-out join evidence from monolingual grammar showing that prosodic, phonological, and morphological phenomena often cross the boundaries defined by complement spell-out (Kratzer & Selkirk 2007; Cheng & Downing 2012; D'Alessandro & Scheer 2015; Bošković 2016). For example, Shlonsky (1994) reports that complementizers in Palestinian Arabic display morphological allomorphy dependent on the identity of the embedded subject, which Bošković (2016) argues is incompatible with TP spell-out. These types of phenomena challenge the basic conception of Phases as derivational units that determine locality domains across modules of grammar, including PF. Accordingly, researchers propose modifications to Phase theory which allow for phonological cycles to match the boundaries of full Phases, while capturing syntactic locality restrictions (D'Alessandro & Scheer 2015; Bošković 2016).

6. Conclusions

This study investigates the idea that intra-sentential code-switching (ICS) is subject to constraints at the syntax-phonology interface, where the output of the syntactic derivation must be split between two language-specific phonological systems. Data from 3 studies on Hebrew-English ICS are reported, showing that bilinguals prefer to switch between separate prosodic phrases, rather than within one. In embedded clauses, bilinguals prefer a complementizer or a bare *wh*-word to match the language of the TP, but have no preference for the language of a prosodically independent *wh*-phrase; in subject DPs, bilinguals prefer the determiner to match rather than mismatch the language of the NP; and finally, bilinguals prefer a pronominal subject to match the language of the verb it is prosodified with, which varies by the Phase status of the surrounding clause. We argue that these data are incompatible with López, Alexiadou & Veenstra (2017)'s proposal that ICS can only take place between spell-out domains determined by standard Phase theory, nor can they be accounted for under MacSwan (1999)'s proposal that ICS cannot take place within a syntactic head. We discuss implications for Phase theory, suggesting that our data are compatible with a theory of full Phase spell-out, and consider potential implementations of our *ICS by prosodic phrase* constraint within the architecture of Distributed Morphology.

References

Arregi, K. & Nevins, A. 2012. *Morphotactics: Basque Auxiliaries and the Structure of Spellout*. Berlin: Springer.

Badiola, L. & Sande, A. 2018. A study of simultaneous bilinguals: Gender assignment in Basque/Spanish mixed determiner phrases. In L. López (ed.). *Code-switching – Experimental Answers to Theoretical Questions: in honor of Kay González-Vilbazo*, 15-38. John Benjamins Publishing Company (Issues in Hispanic and Lusophone Linguistics).
<https://doi.org/10.1075/ihll.19.02bad>

Belazi, H.M., Rubin, E.J. & Toribio, A.J. 1994. Code Switching and X-Bar Theory: The Functional Head Constraint. *Linguistic Inquiry* 25(2): 221-237.

Bennett, R. & Elfner, E. 2019. The Syntax–Prosody Interface. *Annual Review of Linguistics* 5: 151-171.
<https://doi.org/10.1146/annurev-linguistics-011718-012503>

Birdsong, D., Gertken, L.M. & Amengual, M. 2012. Bilingual language profile: An easy-to-use instrument to assess bilingualism. COERLL, University of Texas at Austin. <https://blp.coerll.utexas.edu>

Bolinger, D. 1971. The Nominal in the Progressive. *Linguistic Inquiry* 2(2): 246-250.

Bošković, Ž. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59(1): 1-45.
<https://doi.org/10.1111/j.1467-9582.2005.00118.x>

Bošković, Ž. 2016. What is sent to spell-out is phases, not phasal complements. *Linguistica* 56(1): 25-66.
<https://doi.org/10.4312/linguistica.56.1.25-66>

Bromberger, S. & Halle, M. 1989. Why Phonology Is Different. *Linguistic Inquiry* 20(1): 51-70.

Büring, D. 2013. Syntax, information structure, and prosody. In den Dikken, M. (ed.). *The Cambridge handbook of generative syntax*, 860-896. Cambridge: Cambridge University Press
<https://doi.org/10.1017/CBO9780511804571.029>

Büring, D. & Truckenbrodt, H. 2021. Correspondence between XPs and Phonological Phrases. *Linguistic Inquiry* 52(4): 791-811.
<https://doi.org/10.1162/ling_a_00391>

Bürkner, P.-C. 2017. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* 80(1): 1-28.
<https://doi.org/10.18637/jss.v080.i01>

Caramazza, A. et al. 1973. The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America* 54(2): 421-428.
<https://doi.org/10.1121/1.1913594>

Cheng, L.L. & Downing, L.J. 2012. Prosodic domains do not match spell-out domains. *McGill Working Papers in Linguistics*, 22.

Chomsky, N. 1995. *The Minimalist Program*. MIT Press. <https://www.scirp.org/reference/referencespapers?referenceid=2951613> (Accessed: July 13, 2025).

Chomsky, N. 2000. Minimalist Inquiries: The Framework. In Martin, R., Michaels, D. & Uriagereka, J. (eds.). *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89-155. Cambridge, MA: MIT Press.

Chomsky, N. 2001. Derivation by Phase. In Kenstowicz, M. (ed.). *Ken Hale: A Life in Language*, 1-52. MIT Press.
<https://doi.org/10.7551/mitpress/4056.003.0004>

D'Alessandro, R. & Scheer, T. 2015. Modular PIC. *Linguistic Inquiry* 46(4): 593-624.

Danon, G. 1996. *The syntax of determiners in Hebrew*. Ph.D. dissertation, Tel-Aviv University.

Di Sciullo, A.-M., Muysken, P. & Singh, R. 1986. Government and code-mixing. *Journal of Linguistics* 22(1): 1-24.
<https://doi.org/10.1017/S0022226700010537>

Elfner, E. 2015. Recursion in prosodic phrasing: evidence from Connemara Irish. *Natural Language & Linguistic Theory* 33(4): 1169-1208.
<https://doi.org/10.1007/s11049-014-9281-5>

Embick, D. 2010. *Localism versus Globalism in Morphology and Phonology*. Cambridge, MA: MIT Press.
<https://doi.org/10.7551/mitpress/9780262014229.001.0001>

van Gelderen, E. & MacSwan, J. 2008. Interface conditions and code-switching: Pronouns, lexical DPs, and checking theory. *Lingua* 118(6): 765-776.
<https://doi.org/10.1016/j.lingua.2007.05.003>

González-Vilbazo, K. 2005. *Die Syntax des Code-Switching. Esplugisch: Sprachwechsel an der Deutschen Schule Barcelona*. PhD dissertation, Universität zu Köln [Preprint]. <https://www.academia.edu/741071/Die_Syntax_des_Code_Switching_Esplugisch_Sprachwechsel_an_der_Deutschen_Schule_Barcelona>

González-Vilbazo, K. et al. 2013. Methodological Considerations in Code-Switching Research. *Studies in Hispanic and Lusophone Linguistics* 6(1): 119-138.
<https://doi.org/10.1515/shll-2013-1143>

González-Vilbazo, K. & López, L. 2011. Some properties of light verbs in code-switching. *Lingua* 121(5): 832-850.
<https://doi.org/10.1016/j.lingua.2010.11.011>

González-Vilbazo, K. & López, L. 2012. Little v and parametric variation. *Natural Language & Linguistic Theory* 30(1): 33-77.
<https://doi.org/10.1007/s11049-011-9141-5>

González-Vilbazo, K. & López, L. 2013. Phase switching. Paper presented at *Workshop: Code-switching in the bilingual child: Within and across the clause*, 18. <https://www.academia.edu/7990026/Phase_switching>

Gosselin, L. 2022. Bilinguals have a single computational system but two compartmentalized phonological grammars: Evidence from code-switching. *Glossa: a journal of general linguistics* 6(1): 147.
<https://doi.org/10.16995/glossa.5800>

Grimstad, M.B., Lohndal, T. & Åfarli, T.A. 2014. Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd* 41(2): 213-237.
<https://doi.org/10.7557/12.3413>

Gullberg, M., Indefrey, P. & Muysken, P. 2009. Research techniques for the study of code-switching. In Bullock, B.E. & Toribio, A.J. (eds.). *The Cambridge handbook of linguistic code-switching*, 21-39. Cambridge: Cambridge University Press (Cambridge Handbooks in Language and Linguistics).
<https://doi.org/10.1017/CBO9780511576331.003>

Gumperz, J.J. 1977. The Sociolinguistic Significance of Conversational Code-Switching. *RELC Journal* 8(2): 1-34.
<https://doi.org/10.1177/003368277700800201>

Halle, M. & Marantz, A. 1993. Distributed morphology and the pieces of inflection. In K. Hale & S.J. Keyser (eds.). *The view from building 20*, 111-176. Cambridge, MA: The MIT Press.

Halle, M. & Mohanan, K.P. 1985. Segmental phonology of modern english. *Linguistic Inquiry* 16(1): 57-116.

Joshi, A.K. 1982. Processing of Sentences With Intra-Sentential Code-Switching. *Proceedings of Coling 1982: Ninth International Conference on Computational Linguistics*. *COLING 1982*. <https://aclanthology.org/C82-1023/>

Koronkiewicz, B. 2014. *Pronoun Categorization: Evidence From Spanish/English Code-switching*. PhD dissertation, University of Illinois Chicago. <https://hdl.handle.net/10027/19014>

Koronkiewicz, B. & Ebert, S. 2018. Modality in experimental code-switching research: Aural versus written stimuli. In L. López (ed.). *Code-switching – Experimental Answers to Theoretical Questions: In honor of Kay González-Vilbazo*, 147-176. John Benjamins Publishing Company.
<https://doi.org/10.1075/ihll.19.06kor>

Kratzer, A. & Selkirk, E. 2007. Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24(2-3): 93-135.
<https://doi.org/10.1515/TLR.2007.005>

López, L. 2020. *Bilingual Grammar: Toward an Integrated Model*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/9781108756181>

López, L., Alexiadou, A. & Veenstra, T. 2017. Code-Switching by phase. *Languages* 2(3): 9.
<https://doi.org/10.3390/languages2030009>

MacSwan, J. 1999. *A Minimalist Approach to Intrasentential Code Switching*. Taylor & Francis.

MacSwan, J. 2000. The architecture of the bilingual language faculty: evidence from intrasentential code switching. *Bilingualism: Language and Cognition* 3(1): 37-54.
<https://doi.org/10.1017/S1366728900000122>

MacSwan, J. 2006. Code Switching and Grammatical Theory. In T.K. Bhatia & W.C. Ritchie (eds.). *The Handbook of Bilingualism and Multilingualism*, 282-311. John Wiley & Sons, Ltd.
<https://doi.org/10.1002/9780470756997.ch11>

Mahootian, S. 1993. *A null theory of codeswitching*. PhD dissertation, Northwestern University. <https://www.proquest.com/docview/304089810/abstract/2CF00B1213854DBCPQ/1>

Muysken, P. 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge University Press.

Myers-Scotton, C. 1993. *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Clarendon Press.

Nespor, M. & Vogel, I. 1986. *Prosodic phonology*. Dordrecht: Foris.

Olson, D.J. 2019. Phonological processes across word and language boundaries: Evidence from code-switching. *Journal of Phonetics* 77: 100937. <https://doi.org/10.1016/j.wocn.2019.100937>

Ott, D. 2011. A Note on Free Relative Clauses in the Theory of Phases. *Linguistic Inquiry* 42(1): 183-192.

Pak, M. (2008) *The postsyntactic derivation and its phonological reflexes*. PhD dissertation, University of Pennsylvania. <https://www.researchgate.net/publication/247949599_The_postsyntactic_derivation_and_its_phonological_reflexes>

Parafita Couto, M.C. & Gullberg, M. 2019. Code-switching within the noun phrase: Evidence from three corpora. *International Journal of Bilingualism* 23(2): 695-714. <https://doi.org/10.1177/1367006917729543>

Poplack, S. 1980. Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching 1. *Linguistics* 18: 581-618. <https://doi.org/10.1515/ling.1980.18.7-8.581>

Samek-Lodovici, V. 2005. Prosody–Syntax Interaction in the Expression of Focus. *Natural Language & Linguistic Theory* 23(3): 687-755. <https://doi.org/10.1007/s11049-004-2874-7>

Schuhmann, K.S. & Putnam, M.T. 2021. Relativized Prosodic Domains: A Late-Insertion Account of German Plurals. *Languages* 6(3): 142. <https://doi.org/10.3390/languages6030142>

Selkirk, E. 1986. On derived domains in sentence phonology. *Phonology* 3: 371-405. <https://doi.org/10.1017/S0952675700000695>

Selkirk, E. 1996. The prosodic structure of function words. In Morgan, J.L. and Demuth, K. (eds.). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 187-213. Hillsdale, NJ, US: Lawrence Erlbaum Associates.

Selkirk, E. 2009. On Clause and Intonational Phrase in Japanese: The Syntactic Grounding of Prosodic Constituent Structure: Interfaces of Language Structures. *GENGO KENKYU (Journal of the Linguistic Society of Japan)* 136: 35-73. <https://doi.org/10.11435/gengo.136.0_35>

Selkirk, E. 2011. The Syntax-Phonology Interface. In J. Goldsmith, J. Riggle & A.C.L. Yu (eds.). *The Handbook of Phonological Theory*, 435-484. Blackwell Publishing Ltd. <https://doi.org/10.1002/9781444343069.ch14>

Shlonsky, U. 1994. Semitic clitics. *Geneva Generative Papers* 2: 1-11.

Silber-Varod, V. 2014. Dependencies over prosodic boundary tones in spontaneous spoken Hebrew. In K. Gerdes, E. Hajičová & L. Wanner (eds.). *Dependency Linguistics: Recent advances in linguistic theory using dependency structures*, 207-228. John Benjamins Publishing Company. <https://doi.org/10.1075/la.215.10sil>

Timm, L.A. 1975. Spanish-English Code-Switching: El Porqué y How-Not-To. *Romance Philology* 28(4): 473-482.

Truckenbrodt, H. 1995. *Phonological phrases: their relation to syntax, focus, and prominence*. Massachusetts Institute of Technology. <https://cir.nii.ac.jp/crid/1970023484964388262>

Truckenbrodt, H. 1999. On the Relation between Syntactic Phrases and Phonological Phrases. *Linguistic Inquiry* 30(2): 219-255. <https://doi.org/10.1162/002438999554048>

Vaughan-Evans, A. et al. 2014. Anomalous Transfer of Syntax between Languages. *Journal of Neuroscience* 34(24): 8333-8335. <https://doi.org/10.1523/JNEUROSCI.0665-14.2014>

Wagner, M. 2010. Prosody and recursion in coordinate structures and beyond. *Natural Language & Linguistic Theory* 28(1): 183-237. <https://doi.org/10.1007/s11049-009-9086-0>

Wagner, M. & Watson, D.G. 2010. Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes* 25(7-9): 905-945. <https://doi.org/10.1080/01690961003589492>

Wintner, S. 1998. The affixal nature of the definite article in Hebrew. In *Computational Linguistics in the Netherlands 1997. Selected Papers from the Eighth CLIN Meeting, number 25 in Language and Computers: Studies in Practical Linguistics*, 145-167. Editions Rodopi.

Zehr, J. & Schwarz, F. 2018. PennController for Internet Based Experiments (IBEX). *Open Science Foundation*. <https://doi.org/10.17605/OSF.IO/MD832>