Issues in Basque-Romance Contact: Interrogative Structures
Ricardo Etxepare (CNRS, IKER UMR 5478)

1. BACKGROUND

I-languages in contact:

Most of the work on language contact during the last decennies may be included within the notion of E-language correspondences : case studies in which the borrowing is defined in terms of morphemes, constructions, or other superficial categories belonging in the domain of the immediately observable. The comparison between the languages in contact in those cases is established in terms of similarities (constructional similarities, lexical similarities, or more general notions of similarity at the level of communicative strategies, languages or typological generalizations). The notion of similarity, convergence, or the usual technical lexicon employed in most of the work on language contact is deeply rooted in E-language notions. We will try to show by means of some illustrative cases, that such an approach to contact induced change is not appropriate to (i) describe the relevant change, and (ii) set the basis for an explanatory account of the change.

Take for instance the zero complementizer borrowed by Spanish in contact with English in the United States (Silva-Corbalán, 1994; 2008:216):

(1) a. She told me $\varnothing$ she was delighted with the vase $\varnothing$ she had bought
   b. Me dijo *(que) estaba feliz con el jarrón *(que) había comprado   (Standard Spanish)

(2) Yo creo $\varnothing$ inventaron   el nombre **que** le pusieron   (LA Spanish)
   I think   they.invented the name that CL they.gave
   'I think they invented the name they gave him'

Etxepare (1999), building on work by Ormazabal (1995) and Boskovic (1997), notes that null complementizers in Spanish come in two guises: modal null C, apparent in subjunctive complements (3a), and indicative null C, involved in the syntax of extraction:

(3) a. Les   ruego $\varnothing$ me   devuelva   alguien   el pasaporte
      3PL.DAT I.beg   1S.DAT give.back.SUBJ someone the passport
      "I beg you (formal) that someone gives me back my passport"

   b. El vino que Pedro asegura $\varnothing$ les   recomendaron   a sus amigos
      the wine that Pedro claims   3PL.DAT they.recommended to his friends
    "The wine that Pedro claims his friends were recomended"

Etxepare observes that fronted material inmediately preceding the higher verb makes the null complementizers ungrammatical :

(4) a. *Les   ruego alguien me   devuelva   el pasaporte (cf. Le ruego que usted...)
       3PL.DAT I.beg someone 1S.DAT give.back.SUBJ the passport

   b. *El vino que Pedro asegura a sus amigos les   recomendaron (cf. que a sus amigos...)
      the wine that Pedro claims to his friends 3.pl.dat they.recommended

Etxepare proposes a C-to-V cliticization process happening at PF, and requiring the adjacency of the null C and the V (blocked by the fronted material in (4a,b)). Boskovic and Lasnik (2003) propose a similar analysis for null C in verbal dependents, keeping it separate from the null C in relative clauses. The null C of declarative dependents is an affix with the selectional feature [+V]. Looked at from this perspective, the change in Los Angeles Spanish is a change from null Cs which have a complex morphosyntactic content: [+V$_{morphology}$, +wh] and [+V$_{morphology}$, +modal], to a null C *à l'anglaise* ([+V]). It does not extend to relatives, which do not share the same dependency vis-à-vis the verb.

What is borrowed is not just a zero complementizer $\varnothing$ (an E-language category), but the zero complementizer as it relates to a particular array of features, a given pair [Exponent, (F1, F2,... F$_n$)] which operates selectively over the English configurations in the bilingual mind (see recently Aboh, 2015).

The induction problem in SLA

Poverty of stimulus arguments have been put forward within the context of L2 acquisition. They typically present an scenario in which a particular grammatical regularity, which does not exist in the grammar of the learner, and for which negative evidence is not available, nevertheless ends up being part of the interlanguage of even not very advanced learners. A well known case is the French structure investigated by Dekydspotter and Sprouse (2001):

(5) a. Qui de célèbre fumait au bistro dans les années 70?
    b. Qui fumait de célèbre au bistro dans les années 70?

The meaning of (5a,b) is not identical. A possible answer to the question in (3a) may involve a person who is currently famous but was not famous in the 70s. The discontinuous interrogative can only be answered with a celebrity of the 70s. The structure does not exist in English, to start with, and cannot be taught except by having recourse to negative evidence. Dekydspotter and Sprouse showed that English L2 learners of French, even not very proficient ones, demonstrated a statistically significant difference in the acceptance of the available and unavailable interpretations.

Kanno (1997) (apud Schwartz and Sprouse, 2016).

(6)   Overt Pronoun Constraint

   In languages that permit null arguments, an overt pronoun cannot be interpreted as a bound variable

The OPC provides a straightforward account of the following asymmetry in the interpretive possibilities of null and overt pronouns in Japanese :

(7)   a. Tanaka-san$_i$ wa [$\varnothing_{i/other}$ kaisya   de itiban da to] itte-iru
       Tanaka-Mr TOP   company in best is that saying-is
       'Mr Tanaka says that I/you/he/she…is the best in the company'

     b. Dare$_i$ ga [$\varnothing_{i/other}$ sore o   mita to] itta no ?
        who NOM   that ACC saw that said Q
        'Who said that I/you/he/she… saw that ?'

     c. Tanaka-san$_i$ wa [kare$_{i/other}$ kaisya   de itiban da to]   itte-iru
        Tanaka-Mr TOP he   company in best is that saying-is
        'Mr Tanaka says that he is the best in the company'

     d. Dare ga [kare$_{*i/other}$ sore o   mita to] itta no ?
        who NOM   that ACC saw that said Q
        'Who said that he saw that?'

In (7a), the null subject of the embedded clause can be interpreted referentially, as coreferential with the matrix subject or with an extra-sentential antecedent (*other*). In (7b), the null subject of the embedded clause can be interpreted referentially or as a bound variable. In (7c), *kare* 'he', can be interpreted referentially, again as coreferential to either the subject antecedent or to some other extra-sentential referent. However, while *kare* in (5d) can also be interpreted referentially (that is, as linked to an extra-sentential referent), the bound variable interpretation is impossible. Schwartz and Sprouse qualify this situation as an example of « bankruptcy of the stimulus ». The direct evidence available in the input for the acquirer is summarized in (8):

(8)

	Referential interpretation	*Bound interpretation*
Null pronoun	Attested in the input	Attested in the input
Overt pronoun	Attested in the input	

 This should provide a child exposed to Japanese with evidence for the grammaticality of these three cells, as in (9) :

(9)

	Referential interpretation	*Bound interpretation*
Null pronoun	OK	OK
Overt pronoun	OK	

Natural analogical extension would lead to the conclusion that the overt pronoun can also receive a bound interpretation, as (10) :

(10)

	Referential interpretation	*Bound interpretation*
Null pronoun	OK	OK
Overt pronoun	OK	OK

But this doesn't happen. Children exposed to Japanese know that the interpretation of an overt pronoun as a bound variable is not possible.

(11)

	Referential interpretation	*Bound interpretation*
Null pronoun	OK	OK
Overt pronoun	OK	*OK

This conclusion about the impossible bound-variable interpretation occurs in the absence of any PLD directly relevant to just that interpretation. Japanese children do not receive instruction on this paradigm, and negative evidence is not available.

Kanno showed that the same constraint was operative for adult English native speakers (four semester instruction) learning Japanese, despite the fact that English, not being a null subject language, does not obey that constraint :

(12)   Who$_i$ said that he$_{i/other}$ saw that ?

Kanno used this type of evidence to argue for full access to UG in the course of second language acquisition.

The poverty of stimulus situations in SLA illustrated by cases such as the Overt Pronoun Constraint bear witness to the availability of UG principles in constraining successive steps in bilingual acquisition. It is what the SLA litterature calls "full access" (see White, 2003, for an

overview of the discussion). It also corresponds to a scenario of attainment of a target grammar.

Emerging forms

Contact induced change may result in structures which belong neither to the model language nor to the replica language, another poverty of stimuus scenario.

Emerging forms are apparent in Code-Switching, for instance in the well documented use of light verb constructions in code-switching pairs which are otherwise unattested in either of the languages present in the switch. One such instance are structures like (13), coming from Spanish-German code-switching as studied by Gonzalez-Vilbazo and Lopez (2011) :

(13) *Juan hace* nähen das Hemd
       Juan does sew   the shirt
     « Juan sews the shirt »

Hacer+V cannot be used as a light verb construction in monolingual Spanish. There is no equivalent either in monolingual German. The light verb can only be realized in one language (Spanish), but not the other (German).

This type of asymmetry has also been described for :

Telugu/English contact   (Den Dikken and Rao, 2003)
Marathi/English   (Joshi, 1985)
Hindi/English   (Ritchie and Bhattia, 1996)
Turkish/Dutch   (Boschoeten and Verhoeven, 1985)

How does a structure like (13) emerge ?

Ingredient 1 : A basic clausal structure that divides the verbal phrase into a functional v and a lexical V

(13) $[_{CP}…T [_{vP} v [_{VP} V…]]]$

Ingredient number 2: Spanish verbs have conjugation classes:

(14)   a. *cant-* *"sing"* -ar class 1 -> cantar (ex: cantas = you sing)
          b. *beb-* "drink" -er class 2 -> beber (ex: bebes = you drink)
          c. *viv-* "live" -ir class 3 -> vivir (ex: vives = you live)

Ingredient number 3 : small v possesses a morphological feature *conjclass*, that must be satisfied by a lexical verb.

(15)   *vP
                              vsp[uConj]   VP
                                           ….
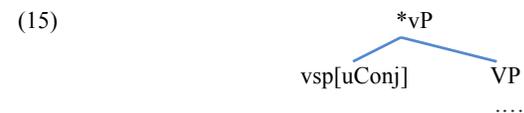
The rest follows from general computational restrictions. German verbs do not have a conjugation class, so they cannot satisfy the conjugation class feature of small v. This means that the German lexical verb does not raise to merge with v. It stays in-situ. In its in-situ position, it adopts an unmarked form corresponding to the infinitive. The absence of V-to-v makes necessary the presence of a last resort empty verb : *hacer*, that can satisfy the conjugation class feature. This phenomenon can be compared to dummy-do insertion in English. The light verb must be the Spanish one. Why ? An alternative in which v would lack a conjugation class feature would be satisfied by ordinary V-to-v movement, so no light verb would be required, and therefore, by economy, it would not be introduced.

The reason why the order of the inner VP corresponds to the Spanish one is because the v-probe corresponds to the Spanish one, and Spanish v does not trigger object shift (the reason why the basic order of German is SOV).

Lopez and Gonzalez-Vilbazo note the following contrast :

(16) a. Juan hizo   bauen ein Haus
         Juan made build   a   house
         'Juan built a house'

      b. Juan hizo   a   Pedro ein Haus bauen
         Juan made PREP Pedro a   house build
         'Juan made Pedro build a house'

What's the difference in the VP between (17a,b) ? Structure of causatives with an inner agent (Guasti, 1992 ; Folli and Harley, 2009) :

(17) a. $[_{vP1}$ EA v^0 $[_{vP2}$ EA v^0 $[…V^0…]]$
      b. $[_{vP1}$ *Juan hiz-*^{0} $[_{vP2}$ *a Pedro* v_{German} $[…bauen…]]$

The second v is a German v. It has no conjugation class feature. It attracts the lexical verb. The syntax of the inner VP corresponds to German (object shift).

Lopez and Gonzalez-Vilbazo : 'The I-language of Esplugish speakers is also the outcome of the interaction of UG with the environment. The resulting I-language may contain a feature – a free standing light verb- that is not present in the input grammars, but is available in the universal pool'.

3

Although the relation of this phenomenon with acquisition is not clear (not clear that the light verb constructions survive outside the code-switching context), it attests to the mediating role of UG AND of the two grammars involved in the contact situation. It is thus an example of Full Transfer, Full Access.

The process of bilingual acquisition and of "partially restructured vernaculars" (Holm, 2003) is thus UG-constrained, part of the UG restrictions being whatever follows from the parameter options already in place in unbalanced bilingual acquisition (the general case).

2. THIRD FACTOR EFFECTS IN LANGUAGE CONTACT

One element that is lacking in ths picture is what Chomsky (2005) calls "third factor considerations". This constitutes a fundamental ingredient of neo-emergentist approaches to language variation. Let me start with Biberauer's idea that acquisition biases AND structure building are driven by a third factor principle, not part of UG, that she describes as *Maximize Minimal Means* (2019). In the context of language, *Maximize Minimal Means* will result in the combined action of the two principles below:

Roberts and Roussou (2003:201):

(18) Feature Economy (FE, intuitive formulation): "Postulate as few FFs as possible, given the PLD"

In the context of L1 acquisition, this may read as follows (Roberts, 2019: 93)

(18') Given a pair of adequate structural representations R, R' for a substring of input text of the PLD S, choose R iff R has n distinct FFs and R' has m>n distinct FFs

(19) Input Generalization (IG): Maximize already postulated features (from Roberts, 2007)

The particular interest of *Maximize Minimal Means* with regard to (18) and (19) is that it is at the same time a principle of structure building and a bias guiding acquisition. Take the following quote regarding Feature Economy (under what conditions must the learner postulate a feature in his grammar?):

"...an acquirer who does not pick up on a systematic departure from Saussurean arbitrariness in the input will not pose the 'F present?' question, with the result that the initial NO is a default which the comparatively oriented linguist can juxtapose with the initial YES, the answer that necessarily results when some form of triggering data [Agreement, doubling and expletives, systematic null exponence, multifunctionality, and so on] leads to the question beng posed. The initial NO then respects both FE and IG...The initial YES violates FE, but...respects IG as the newly identified F is assumed to be present on all heads in the relevant domain..."

If the postulated ALL does not help account for the domain distribution of the trigger, then further subdivisions of F must be entertained (SOME).

(20) NONE>ALL>SOME learning path (Biberauer and Roberts, 2016, 2017)

F present?

NO   YES: All heads?

YES   NO: which subset of heads?
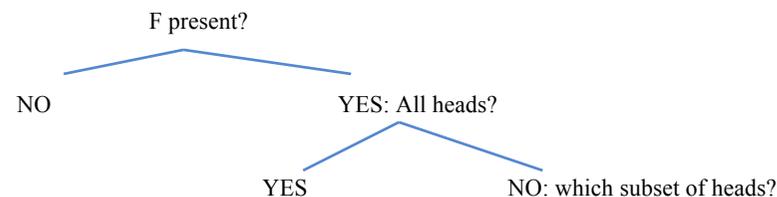
Imagine a 2L1 situation, one where a child is trying to make sense of the varying input in his/her bilingual environment. The child must attend to possible alternative ways to parse the PLD, and I will assume that this parse is as conservative as possible. That is, if Maximize Minmal Means is as much a learning bias as it is a principle of structure building, we may expect something like (20):

(20) Avoid constructing more than one grammar (favor overlap)

Feature Economy (revised for bilingual acquisition)

(21)   Given an adequate structural representation R for a substring of input text of the PLD S, keep postulating R for new incoming input of a language S'. Add R' only if forced.

(21) is intended as a rendition of the intuitive (1) in the context of the construction of two grammatical systems in parallel (bilingual acquisition). (21) should be coupled with Input Generalisation:

(22)   Maximize already postulated features

As this is a principle that governs both learning and structure building, it must apply across the border (it must encompass both languages).

In a context like this, in which input is going to be more varied than in L1 acquisition, we can expect initial NOs to be less frequent. This basically means that language contact (bilingual acquisition) is a powerful engine of language change. The idea is in line with other recent approaches to language contact in the context of creoles (De Graff, 2003 ; Aboh, 2015), now understood as extreme cases of language restructuring under contact (with the Language Bioprogram hypothesis –Bickerton 1984- largely discredited).

In this handout I examine two instances of « partial restructuring » that are led by the third factor MMM, as spelled out in (21) and (22). The first case concerns Basque-Spanish contact, in which the Basque setting (the hypothesis for the existence of a given Feature F in Basque) leads to a somewhat narrower Spanish grammar, as the Spanish setting and the Basque one are in a subset relation. The second case concerns work on Basque-French contact made by Duguine and Irurtzun (2014), that I re-examine on the basis of (21)-(22). In the context of French-Basque contact, with French as an initial setting, it will bring important changes (a superset grammar that combines the Basque and the French setting).

3. PARTIAL INTERROGATIVES AND ADJACENCY IN THE SPANISH OF THE BASQUE COUNTRY (DOLD, ETXEPARE AND KAISER, IN PROGRESS.

Both Basque and Spanish are wh-fronting languages which typically display wh-phrase-verb adjacency, as the subject inverts in wh-questions (Bosque and Gutierrez-Rexach, 2011: 449; Etxepare and Ortiz de Urbina, 2003: 495ff; Aldai, 2011):

(22)　　a. Qué compra Pedro?
　　　　　what buys　Pedro
　　　　　'What does Pedro buy?'

　　　　　b. Zer　erosten　　　du Pellok?
　　　　　　what buy.NOM.LOC AUX Pello.ERG
　　　　　'What does Pello buy?'

(23)　a. *Qué Pedro compra?
　　　　　what Pedro buys

　　　　b. *Zer Pellok　　erosten　　　　du?
　　　　　　what Pello.ERG BUY.NOM.LOC AUX

Torrego (1984), among others, showed that in Spanish, non-argumental wh-phrases don't have to display subject inversion obligatorily, in particular *porqué* 'why':

(24) Porqué (Juan) quiere salir (Juan) antes　que los demás?
　　　Why　Juan　wants leave Juan before than the others
　　　'Why does Juan want to leave before the others?'

For causal wh-words specifically, Basque also allows intervening subjects:

(25) Zergatik (Jonek)　besteek　　baino lehenago atera nahi　du (Jonek)?
　　　why　　Jon.ERG others.ERG than　before　　leave want AUX Jon.ERG
　　　'Why does Jon want to leave before the others?'

Basque and Spanish part ways when complex wh-phrases are at stake. In that case, Spanish does not require wh-phrase-verb adjacency (Ordoñez, 1998), but Basque still does:

(26)　A　cuál　de las chicas (tu hermana) había visitado ya　　(tu hermana) antes?
　　　PREP which of the girls　your sister　had　visited　already your sister　before
　　　'Which of the girls your sister had already visited before?'

Wh-phrase-verb adjacency is still required in Basque in those cases:

(27) Neska horietako zein (*zure arrebak)　bisitatua zuen aurretik (zure arrebak)?
　　　Girl　that.of　which your sister.ERG visited　had　before　your sister.ERG
　　　'Which of the girls your sister had already visited before?'

3.1. Dold (2018)

There is a significant statistical difference in the acceptability of non-adjacent wh-phrase-verb configurations among Basque/Spanish bilinguals with regard to monolingual Spanish speakers. This statistical difference affects complex wh-words, not *why*-questions.

Dold has interviewed 114 subjects in the Basque Country and Spain, divided according to their bilingual/monolingual status, and within the bilingual group, according to the age of exposure to Basque and Spanish. All bilingual speakers interviewed are fluent speakers of Basque.

(28)

Participants

Speaker groups		#
Simultaneous Learners:		
2L1		35
Sequential Learners:		
BSce	= Basque - Spanish, child early	3
BScl	= Basque - Spanish, child late	21
SBce	= Spanish - Basque, child early	8
SBcl	= Spanish - Basque, child late	1
Monolingual Speakers:		
MonoPV	= monolinguals, Basque Country	23
MonoE	= monolinguals, rest of Spain	23
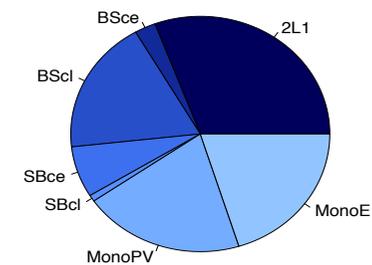
Age steps for the sequential learners after Meisel (2001) and Montrul (2008)

Differences between bilinguals and monolinguals regarding sentences like (29):

(29) A cual de los enfermos el médico examinó durante más de media hora?
      PREP which of the.PL sick.PL the doctor examined for more than half an hour?
      'Which of the sick people did by the doctor examine for more than half an hour?'

Significantly lower acceptance rate by all bilingual speakers compared to monolinguals (p<0.01)

Significant higher acceptance rate by monolinguals from outside the Basque Country, as compared to monolinguals from the Basque Country.

Significantly lower acceptance rate by 2L1 speakers as compared to the sequential bilinguals (p< 0.03)

No clear differences between the sequential learners, probably also due to the small number of participants in three of the four groups.

One of the possible explanations of the different acceptability patterns of bilingual and monolingual populations, and of the early bilinguals as compared to late bilinguals is that this reflects the effect of the L1 Basque on the L2 Spanish.

How should we model this effect? From an E-language perspective, the learner has evidence that Spanish admits sequences such as (29) : they are present (to which extent is a different matter) in the input. Somehow this evidence is neglected.

3.2. An excursus in Basque word order

Basque is a language with basic subject-object-verb (SOV) order, but with a relatively free word order. However, SOV is statistically the most frequent order and pragmatically the more neutral (Rijk 1969, Ortiz de Urbina 1995, 2003, Elordieta 2001), as in (30).

(30) Atzo Jonek Mireni liburua eman zion.
     Yesterday John.ERG Mary.DAT book.the.ABS gave AUX
     'Yesterday, John gave Mary the book.'

Basque has postpositions, possessors precede possessees, relative clauses precede their antecedents and embedded clauses precede the complementizer. Basque also shows unexpected noun-adjective order:

(31) Jonek datorren astean liburu garesti bat erosiko duela esan du
      John.ERG come.REL week-in book expensive one buy.FUT AUX.COMP said AUX
      'John said that he will buy an expensive book next week.'

Deviations from the neutral sentential order are possible, as far as the immediate preverbal element is focalized (and assigned contrastive stress) (Etxepare & Ortiz de Urbina 2003, an observation due originally to Altube 1929, and called in the Basque grammatical tradition "Altube's generalization"). Note that "preverbal" here means preceding the "aspect bearing lexical verb":

(32) a. Atzo JONEK eman zion Mireni liburua.
       yesterday John.ERG give.PARTC AUX Mary.DAT book.the.ABS

     'Yesterday JON gave the book to Miren'

   b. Atzo MIRENI eman zion Jonek liburua
      yesterday Mary.DAT gave. PARTC AUX John.ERG book.the.ABS

     'Yesterday Jon gave the book to MIREN'

   c. Atzo LIBURUA eman zion Jonek Mireni.
      yesterday book.the.ABS gave. PARTC AUX John.ERG Mary.DAT

     'Yesterday John gave Mary the BOOK.'

3.3. Where are wh-phrases in Basque?

The received view is that Basque focus and wh-questions are displaced to a dedicated position in the left periphery of the clause. In the traditional view, this was the specifier of Comp (Ortiz de Urbina 1989):

(33)   a. Nork erosi du liburua?
          who.ERG bought AUX book.the
          'Who bought the book?'

        b. [CP Nork [erosi+du C [IP (nork) (erosi+du) [[vP (nork) liburua (erosi)]]]

(34) a. JONEK erosi du liburua
       Jon.ERG bought AUX book.the
       'It is Jon who bought the book.'

     b. [CP Jonek [erosi+du C [IP (Jonek) (erosi+du) [[vP (nork) liburua (erosi)]]]

The unitary treatment of wh- and focus-syntax makes sense in Basque. We can put forward several arguments in favour of this approach or some related approach. Note that both precede negation in Basque (an IP-external category). Consider (35):

(35) Ez   da   etorri
    NEG AUX come
    'He/she did not come'

As observed by Laka (1990), negation can be shown to dominate the inflectional phrase in Basque. One structural configuration that shows this is ellipsis:

(36) a.Jonek   liburua   leitu du, baina Mirenek   ez.
    Jon.ERG book.the read AUX but   Miren.ERG NEG
    'Jon read the book but Miren didn't'

    b.Jonek liburua leitu du baina [NegP Mirenek ez [IP ~~leitu du~~]]

The presence of negation licenses IP-ellipsis in Basque, showing that it dominates the Inflection Phrase. The obligatory position of wh-phrases in Basque is preceding negation, not following it:

(37) a. Nork ez   du   liburua   leitu?
    Who NEG AUX book.the read
    'Who didn't read the book?'

    b. Zer ez   du   Jonek   leitu?
    What NEG AUX Jon.ERG read
    'What didn't Jon read?'

The same goes for focus (but see later):

(38) a. JONEK ez   du   liburua   leitu
    Jon.ERG NEG AUX book.the read.PARTC
    "It is JON who didn't read the book"

    b. LIBURUA ez   du   Jonek   leitu
    book.the   NEG AUX Jon.ERG read.PARTC
    "John didn't read THE BOOK"

If negation dominates the inflectional phrase and the wh-phrases precede negation, we must conclude that wh-phrases occupy a syntactic position above Negation Phrase.

3.4. Shared syntactic properties of foci and wh-phrases in Basque

Basque foci and wh-phrases share a number of parallel properties. We note: (i) the possibility of long distance extraction; (ii) cyclicity; (iii) mutual intervention; (iv) massive pied-piping; and (v) sensitivity to islands.

3.5.1. Long distance extraction

(39) a.Nor$_i$ esan dute   [uste dutela [_$_i$ etorri dela]]?
    Who said AUX.3PL think AUX.COMP   come AUX.COMP
    'Who did they say that they think has come?'

    b. JON$_i$ esan dute   [uste dutela   [_$_i$ etorri dela]]
    Jon   said AUX.3PL think AUX.COMP   come AUX.COMP
    'It is JON who they said that they think has come'

3.5.2. Cyclicity

(40) a.Nor$_i$ esan du Mikelek [uste duela   Jonek [_$_i$ etorriko   dela]]?
    Who said AUX Mikel.ERG think AUX.COMP Jon.ERG   come.PROSP AUX.COMP
    'Who did Mikel say that Jon thinks will come?'

    b. *Nor$_i$ esan du   Mikelek [Jonek uste   duela   [_$_i$ etorriko dela]]
    who   said AUX Mikel.ERG Jon.ERG think AUX.COMP come.FUT AUX.COMP
    'Who did Mikel say that Jon thinks will come?'

(41) a. JON$_i$ esan du Mikelek [uste duela   Aitorrek [_$_i$ etorriko dela]]
    Jon said AUX Mikel.ERG think AUX.COMP Aitor.ERG   come.FUT AUX.COMP
    'It is JON that Mikel said that Aitor thinks will come.'

    b. *JON$_i$ esan du Mikelek [Aitorrek uste duela [_$_i$ etorriko dela]]

3.5.3. Mutual intervention in long distance movement

(42) *Zer$_i$ uste du Mikelek   [ETXEAN _$_i$ aurkitu duela   Jonek]?
    what think AUX Mikel.ERG home.LOC   found   AUX.COMP Jon.ERG
    *'*What does Mikel think that it is at HOME that John has found?'*

(43) *Zer$_i$ galdetu du   Mikelek [non _$_i$ aurkitu duen   Jonek]?
    What asked   AUX Mikel.ERG where found AUX.COMP Jon.ERG
    *'*What did Mikel ask where Jon had found?'*

3.5.4. Massive pied-piping

Pied-piping (Ortiz de Urbina 1993, Etxepare 1997, Arregi 2003, Etxepare & Ortiz de Urbina 2003)

(44) a.Nor$_i$ esan duzu [$_{CP}$ uste dutela [$_{CP}$ _$_i$ etorri dela]]?
    Who said AUX.2SG think AUX.COMP come AUX.COMP
    'Who did you say they think has come?'

   b. [$_{CP}$ Nor etorri dela]$_i$ esan duzu [$_{CP}$ uste dutela _$_i$]?
     Who come AUX.COMP say AUX.2SG think AUX-COMP
    'Who did you say they think has come?'

(45) a.JON$_i$ esan dute [$_{CP}$ uste dutela [_$_i$ etorri dela]].
    JON said AUX3PL think AUX.3SG.COMP come AUX.COMP
    'It is JON who they said they think has come.'

   b. [$_{CP}$ JON etorri dela]$_i$ esan dute [CP uste dutela _$_i$].
     JON come AUX.COMP said AUX think AUX.COMP
    'It is JON who they said they think has come.'

3.5.5. Sensitivity to islands

Adjunct islands as an illustrative case:

(46) a.*Nor$_i$ aldegin duzu [_$_i$ etorri delako]?
    Who left AUX.2SG come AUX.COMP.CAUSAL
    'Who did you leave because he/she came?'

   b.*JON aldegin duzu [_$_i$ etorri delako]
    Jon left AUX.2SG come AUX.COMP.CAUSAL
    'It is JON who you left because he/she came'

All strong islands can be circumvented through pied-piping, with both wh-questions and foci (Arregi 2003, Etxepare & Ortiz de Urbina 2003). Note that the pied-piped clause must be adjacent to the verbal complex (see recently Duguine, 2020):

(47) a. [Nor etorri delako] (*Jonek) aldegin du Jonek?
    Who come AUX.COMP.CAUSAL (Jon.erg) left AUX Jon.ERG
    'Who is the x such that Jon left because x came?"

   b. [JON etorri delako] (*Mirenek) aldegin du Mirenek
    Jon come AUX.COMP.CAUSAL (Miren.ERG) left AUX Miren.ERG
    'Miren left because JON came.'

3.6. What are interrogative clause-types made of?

Focus constructions and interrogative constructions share important semantic similarities, that have syntactic structural correlates of the type we have seen. A common semantic treatment of partial questions is that they denote sets of alternative propositions, understood to be the set of all the potential answers to the question. So the meaning of (48a), in the discourse universe in which (48b) are the possible alternatives Peter may have entertained, is the set of propositions in (48c). The abstract denotation of the question in (48a) can be modeled as in (48d):

(48) a. What did Peter eat today?
   b.Alternatives in the discourse universe: {pasta, vegetables, lamb}
   c.{{Peter ate pasta}, {Peter ate vegetables}, {Peter ate lamb} }
   d.λp [Ex $\notin$ human. p=Peter ate x]

Foci present a very similar semantics, where alternatives to the focused element constitute a basic element of their denotation. In the alternatives approach to foci (Rooth, 1992), focused elements have both an ordinary denotation, and a special one, formed by all the alternatives of the same semantic type that may have occupied the place of the focus in the sentence (the alternatives):

(49) a. JOHN came
   b. Ordinary denotation (assertion): John came
   c. Focus semantic value: λp [Ex $\in$ human. p=x came]

The answer to a question is the assertion component of a focused sentence.

On the other hand, interrogative clauses have something that focused sentences such as (49) don't have: they constitute requests for information. They possess a particular force, which assertions lack, and which must be somehow read from the clausal structure. Following Rizzi (1997) let us call Force that part of the clausal structure which is related to the typing of a sentence as interrogative; and Focus, that part of the interrogative clause which is related to the basic alternatives introduced by questions and focused sentences. The left periphery of the clause is endowed with syntactic positions which allow us to identify those two aspects of the meaning of questions.

(50) Force…Focus…

It is quite plausible that other aspects of the meaning of interrogatives lead us to enrich (33). We are thinking particularly on recent work on the syntax of Q-particles (Reinhart, 1995; Hagstrom, 2008; Cable, 2010; Slade, 2012; Chernova, 2014, among others). In any case, those approaches must also assume at least this much.

3.6.1. On the distribution of wh-phrases

We may wonder whether wh-phrases cross-linguistically relate in the same way to Force and Focus. Here's what I mean by this. Take Spanish, a Romance language in contact with Basque. It has a wh-word paradigm (51) whose use is limited to contexts that we would define as +wh, or in other words, as involving clausal typing.

(52) a. *Quién* 'who'
      b. *Qué* 'what'
      c. *Cuál* 'which'
      d. *Dónde* 'where'
      e. *Cuándo* 'when'
      f. *Cómo* 'how'

Those wh-phrases only occur in (free) relatives, and in questions.

(53) a. Quién ha venido?
         Who has come
      'Who came?'

      b. Quien venga   antes se llevará el premio.
         who come.SUBJ before CL take.FUT the prize
      'The one who comes earlier will get the prize'

Basque wh-phrases do not occur in relatives, not even in free ones. Basque ordinary relatives have no overt wh-word operator:

(54) a. Etorri den   gizona   (headed relative)
         come   AUX.REL man.DET
      'The man who came'

      b. Etorri dena   (free relative)
         come AUX.REL.DET
      'Who came'

Then, Basque wh-phrases occur outside wh-constructions, constructions that have no interrogative force. Take the following cases, from Etxepare (2002; Etxepare, forthcoming, for an analysis in terms of an underlying distributive operator, cf. Chinese *dou*):

(55) a. Nork   bere lekua   hartu du
         who.ERG his place.the taken AUX
      'Everyone took his/her place'

      b. Zein bere etxean   gelditzea onena da
         which his   house.LOC stay.nom.the best.the is
      'The best is that everyone stays home'

(55a,b) differ minimally from questions. They become interrogative constructions if the wh-word immediately precedes the verbal complex:

(56) Nork   hartu du bere lekua?
      who.ERG taken AUX his/herplace
      'Who has taken his/her place?'

In other words, converting (56a,b) into a question requires focalizing the wh-word. Also *multiple wh-partitive constructions* (Liptak, 2001):

(57) Nor oinez   nor zaldiz iritsi   ginen
      who walk.by who horse.by arrive.PARTC AUX.PAST
      "We arrived some walking some on horse"

If we look at the relation that wh-phrases establish with the two basic elements of the interrogative left periphery, Force and Focus, we could say that Basque wh-phrases, unlike Spanish ones, do not seem to be related to the expression of Force. They do not determine clausal type, they are absent from wh-structures like relatives, and they can be used outside interrogative constructions.

3.6.2. Wh-word based quantifiers

Basque has a relatively rich system of quantificational expressions based on "indeterminate pronouns" (Kuroda, 1968): indefinite bases formally identical to wh-words. Consider the following table:

(58)

	Wh-words	Existential "Someone"	Polarity "Anyone"	Free Choice 1	Free Choice 2
WHO	nor	nor-**bait**	**i**-nor	**edo**-nor	nor-**nahi**
WHICH	zein	zen-**bait**	(e-zein)	**edo**-zein	zein-**nahi**
WHAT	zer	zer-**bait**	e-zer	**edo**-zer	zer-**nahi**
WHERE	non	non-**bait**	**i**-non	**edo**-non	non-**nahi**
WHEN	noiz	noiz-**bait**	**i**-noiz	**edo**-noiz	noiz-**nahi**
HOW	nola	nola-**bait**	**i**-nola	**edo**-nola	nola-**nahi**

The prefixes:

i/e- for polarity indefinites has been claimed to be a reduced form of negation *ez* "not, no" (Agud and Tovar, 1991, apud Azkue). The combination of **i-/e-** and the wh-word gives rise to NPIs (see Etxepare, 2003):

(59) a. *Ez   da inor/*norbait   etorri*   (Negation)
        NEG AUX anyone/someone come
    "Noone came"

   b. *Inor/*norbait   baino lehenago iritsi   da*   (Comparatives)
      anyone/someone than   earlier   arrived AUX
    "He/she got here before anyone else"

   c. *Inor etortzen bada,…*   (Conditional Protases)
      anyone come if.is
    "If anyone comes"

   d. *Inor   etorri al da?*   (Interrogatives)
      anyone come INT is
    "Did anyone come?"

edo- is identical to the disjunction "or" as in (60):

(60) *Xabier edo Miren*
      Xabier or Miren

In combination with wh-words it yields free-choice indefinites, which are licensed in non-veridical, non-episodic contexts (see Giannakidou, 2001):

(61)   a. *Edozeinek egin dezake*
         or.which do   can
    "Anyone can do it"

    b. *Hori edozeinek egiten du gaur egunean*
       that anyone.ERG do.HAB has today day.in
    "These days anyone does that"

     c. **Edozeinek   egin zuen*
         or.which.ERG done he.had
      "*Anyone did it"

The suffixes:

-nahi is the lexical verb "want". The free-choice series based on *want* is derived from the grammaticalization of an impersonal free relative (cf. Latin *qui-libet*, Spanish *quienquiera* or Roumanian *cine-va* (Haspelmath, 1997; Aloni, 2016):

(62) a. *Nor   nahi den,   eta nolakoa nahi den   zarela*
        WHO.ABS want is.REL, and how   want is.REL you.are.COMP

     "Whoever it is, and however it is that you are"   (Axular, 1643)

    b. *Entzun,   nornahi   zarela   ere*   (Igerabide, 2011)
       listen.IMP who.want you.are.COMP even

     "Listen, whoever you are"

-bait in the existential pronouns is identical to the complementizer *–bait* in finite subordination (Azkue, 1925; Lafon, 1944, 1966; Trask, 1997; Michelena, 1970; Etxepare, 2001; Rebuschi, 2003). I have argued (Etxepare, forthcoming), that those indefinite pronouns come from correlative protases:

(63) Nor ere bait-a   -> Norbait
     who even COMP-is   who.COMP
     "Who (ever) it is"   "Someone"

The paradigm in (58) invites the conclusion that Basque wh-pronouns provide the nominal base (the domain of quantification) for complex quantificational structures. Basque is thus typologically similar in this regard to other languages in which polarity items, certain quantifiers and interrogative pronouns share a common core (see Haspelmath, 1997; Bhatt, 2001, for a typological survey and discussion). How should we characterize wh-phrases in this context?

Let us start with the previous observation regarding the deep similarities between focus and wh-questions. Let us adopt the following idea, from Beck (2006):

(64)   The denotation of wh-phrases is their focus semantic value. Wh-phrases do not have an ordinary denotation, they only denote a set of alternatives.

Semantics of *zer* 'what':

(65)   a. Normal semantics [[zer]] = undefined
      b. Focus semantics [[zer$_{Focus}$]] = {x:x $\in$ inanimate}

Semantics of *nor* 'who':

(66) a. Normal semantics [[nor]] = undefined
      b. Focus semantics [[nor$_{Focus}$]] = {x:x ∈ animate}

The alternative semantics of wh-phrases provides the basis for the logical operations represented in the table (58). Take the suffix *–bait*, used to create the existential *someone/something* series in Basque. A suffix like *–bait* may be taken to introduce a choice function (Reinhart, 1995; 2006) over the alternatives contributed by the wh-word. Choice functions are functions that take a set as their argument and yield a member of that set as its value. If the alternatives contributed by *nor* in a given discourse universe are, say, {Jon, Miren, Peru} the application of the choice function f to this set will yield a particular member of that set:

(67)   f (Jon, Miren, Peru)= Miren

Consider next the free choice pronoun *edonor*, formed by combining the disjunction *edo* "or" and the wh-word. For these cases we will adopt Jayaselaan's idea (1999) that disjunction behaves as a logical operator inducing generalized disjunction. That is, if *nor* "who" in a given discourse context, represents the set of alternatives {Jon, Miren, Peru}, *edonor* yields the set {Jon or Miren or Peru}.

(68) F$_{DISJ}$ (Jon, Miren, Peru) = (Jon or Miren or Peru)

And so on. In languages like Basque, the wh-word providing the alternative basis for quantification can be isolated morphologically and syntactically, as shown by other complex quantifications involving overt binomial conjunctions:

(69) Nor edo nor
        who or   who
        'Someone or other'

In languages like Spanish, the irregular morphological shape of the wh-word series, and the absence of a regular paradigm of the sort in (46) make it difficult to isolate the alternatives component of the interrogative phrase as an independent element.

3.6.3. Massive pied-piping

One of the noteworthy aspects of Basque wh-constructions concerns the possibility of massive pied-piping, a fact we mentioned previously. We repeat the relevant examples below:

(70) a. [Nor etorri delako]   (*Jonek) aldegin du   Jonek?
          who come AUX.COMP.CAUSAL (Jon.ERG) left   AUX Jon.ERG
          "Who is such that Jon left because he/she came?"

      b. [JON etorri delako]   (*Mirenek) aldegin du   Mirenek
          Jon come AUX.COMP.CAUSAL (Miren.erg) left   AUX Miren.ERG
          'Miren left because JON came'

In massive pied-piping, island structures seem to be carried over by the wh-operator inside the island. Pied-piping is obviously an ordinary thing when it comes to wh-questions or focus, as in Spanish:

(71) Acerca de qué habéis   hablado?
      About of what have.2pl talked
      'What did you talk about?'

Pied-piping is a general problem for those approaches of wh-movement that postulate a checking or agreement relation between the wh-word and C, on the basis of a shared wh-feature. The reason is that all pied-piping configurations are configurations in which the wh-word is not in the right structural relation with C to allow agreement or checking. The wh-word does not head the structure. A different issue has to do with the range of variation that the phenomenon allows: pied-piping is an ordinary fact of wh-movement, but massive pied-piping is not. The latter is only available to a subset of those languages that allow pied-piping. We can define massive pied piping as the type of pied piping that allows avoiding strong islands.

Cable (2010) proposes a general approach to pied piping that combines well with the alternative semantics analysis we provided of wh-phrases. Cable finds morphological evidence in languages like Tlingit, Sinhala or Japanese, that in fact there is no relation whatsoever between the higher interrogative feature (call it Force) and the wh-word inside the pied-piped structure. The relation between Force and the constituent embedding the wh-word is mediated by what he calls a Q particle, as in Tlingit below:

(72) a.Waa sa sh tudinookw i   eesh?
        How Q he feels   your father
        'How does your father feel?'

      b. Aadoo yaagu sa ysiteen?
        whose boat   Q you.saw.it
        'Whose boat did you see?'

c. [[waa yateeyi CP] shaax'w saani NP] sa ash   koodlenxaa?
   How they.are.REL girls   type   Q they.are tempting.him
   'What kind of girls are tempting him?' ('Girls that are how are tempting him?')

In Tlingit, all wh-phrases are headed by a particle Q, which attaches to the left of the entire wh-phrase. In cases where massive pied-piping occurs, the particle is merged with the entire island. Cable's analysis of these and similar facts in other languages is that there is no direct relation between Force and wh-s generally, but the relevant relation is established between an interrogative feature in Force and the particle. The particle is, according to Cable, a choice funtion variable f, one that when applied to a set denoting element, yields an entity of that set as its value. The alternatives in (73a,b,c) correspond to the possible values of the variable contributed by the wh-word:

(73) a. {manners x of feeling}
      b. {boats of x}
      c. {Girls who are x-way}

The semantic import of an interrogative would thus be something like: give me all the functions f such that when they apply to the alternatives encoded by the wh-word yield the true propositions that constitute the answer to the question.

3.6.4. Merging Q and locality restrictions in pied-piping

The thing is, how can we account within this view for the difference between the languages that have massive pied-piping and those that don't? Cable suggests that the difference relates to the way in which the particle merges with the constituent containing the wh-word. In some languages, this particle merges as a head with that constituent (74a); in others it merges as an adjunct (74b):

(74) a. [$_{QP}$ Q [$_{XP}$…wh…]]
      b. [$_{XP}$ Q [$_{XP}$…wh…]]

The structure in (74a) is one that allows agreement between the Q and something in its agreement domain. For languages represented in (74a), it is conceivable that an agreement relation is established between the particle and the wh-word, and this restricts the domain in which pied-piping is possible. Spanish for instance:

(75) a. No   sabía [$_{CP}$ qué lata$_i$ habíais tardado tanto   en abrir _$_i$].
         Neg knew.1sg what can had.2pl lasted   so.much in opening
         'I didn't know which can you required so much time to open'

b. *No sabía [$_{CP}$ [en abrir qué lata]$_i$ habíais tardado tanto _$_i$]
     neg knew.1sg in opening what can had.2pl lasted so.much
     'I didn't know which can you required so much time to open'

c. No sabía [$_{CP}$ que [en abrir esa lata]$_i$ habíais tardado tanto _$_i$].
    neg knew.1sg that in opening that can had.2pl lasted   so.much
    'I didn't know that in opening that can, you required so much time'

Pied-piping of an infinitive is not possible in Spanish, as shown in (75b). Note that there is nothing against the displacement of that constituent (75c), as there is no wh-word inside. One possibility is that agreement between Q and the wh-word, cannot be established across lexical categories such as infinitivals, which block agreement in Spanish.

In languages like Basque, we could postulate that Q merges as an adjunct to the phrase containing the wh-word, and that agreement is not an option under such a configuration. In Spanish, Q is merged as a head with the phrase containing the wh-word, and agreement restricts the domain in which pied-piping is possible. This agreement is reflected in the form of wh-phrases, as the *qu*-part (*qu-ien*, *cu-ando*, *qu-é*…)

3.6.5. Targetting Force

We have said that Basque wh-phrases, and in fact wh-phrases generally, do not directly target Force. Is there any evidence that Q targets Force for a language like Basque? One source of evidence for this possibility is provided by the existence of intervention effects. In Basque, we cannot have a wh-word within the scope of negation, despite the fact that this is a possibility for focused constituents and moved wh-phrases:

(76) a. Liburua ez du   XABIERREK ekarri.
         book.the neg AUX Xabier.ERG   brought
         'It is not Xabier who brought the book.'

      b. *Liburua ez   du nork   ekarri?
          book.the NEG AUX who.ERG brought
         'Who is it that didn't bring the book?'

      c. Nork   ez   du liburua ekarri?
         Who.ERG NEG AUX book.the bring.PARTC
         "Who didn't bring the book?"

Cable, following Beck (2006), suggests that the deviant status of (76b) and other analogous cases, follow from the fact that some quantificational element stands in the way of the wh-word (the variable contributing alternatives) and the Q-particle that is supposed to bind it. The general offending configuration is (77):

(77) Q…Neg (or other quantifiers)…wh-word

Dold (2018) observes that if this is correct, then (76b) constitutes evidence that the Q particle must have undergone movement to some position in the left periphery. Assuming that the phrase containing the wh-word raises to focus in Basque, the other plausible target for this displacement is Force.

For languages like Spanish, in which Q is a head, we can assume that it is the Q-particle and its complement (the constituent that contains the wh-word) that raise as a whole to a position in the Agreeing domain of Force (on the basis of the different behavior of simple wh-words like *qué* and complex ones in adjacency, he proposes they raise to a position encoding Contrast (see also Vermeulen, 2012).

3.7. Back to Basque Spanish questions

3.7.1. Distinguishing Basque from Spanish

We could tie those differences to the different derivations that wh-questions undergo in Basque and Spanish. In Basque, only the Q-particle raises to Force, a high left periphery head, whereas wh-phrases only move to Focus, a position that in Basque is adjacent to the verbal complex. In Spanish, the Q-particle and the wh-phrase are merged together and must move as a single constituent. The first step in this movement will be the focus projection, which in Spanish is also a lower position.

The second step in the movement would be raising the Q-particle and the wh-phrase to Force, or to a projection higher than Foc. The final landing site of this movement opens the possibility of inserting topics in between Force and the focus head, that precedes IP. That is impossible in Basque (see 78b), because the wh-word is in the focus position, and nothing can intervene between the focus and the verbal phrase.

(78) a. $[_{ForceP}$ [Q-WhPh] Force^0 $[_{TopP}$ [DP] Top^0 $[_{FocP}$ (Q-WhPh) Foc^0 $[_{IP}$ ...]]]]
      b. $[_{ForceP}$ [Q] Force^0 $[_{TopP}$ [DP] Top^0 $[_{FocP}$ [WhPh] Foc^0 $[_{IP}$ ...]]]]

If the relevant position is not available to simple wh-words like *qué* "what" in Spanish, then we expect them to require adjacency (they will only raise to focus):

(79) *Qué el médico te ha dicho...
       what the doctor cl has said
   "What did the doctore tell you..."

3.7.2. What about why?

Zergatik 'why' occupies a very particular position in the wh-paradigm of Basque. It does not allow complex quantification. Compare:

(80) a. *Nonbait* 'somewhere' (< *non* 'where' + *bait*)
      b. *Zergatikbait* 'for some reason' (*zergatik* 'why' + *bait*)
      c. *Zerbaitengatik* 'for some reason' (*zerbait* 'something' + *-engatik* 'because of')

Zergatik can be used as a relative pronoun/complementizer in colloquial registers:

(81) A: Zergatik joan zara?
         Why   left   you.have
      'Why did you leave'?

      B: Zergatik aspertzen nintzen
         Because bore.IMP I.was
      'Because I was getting bored'

And it is compatible with preverbal focus:

(82) Zergatik INORK ere ez   du liburu hori erosi?
      Why   anyone.ERG even NEG AUX book that bought
   'Why didn't ANYONE buy that book?

Compare:

(82) *Nori   INORK   ERE ez dio liburu hori erosi?
       Who.DAT anyone.ERG even neg aux book that bought
   'Who didn't anyone buy this book to?'

The natural thing to conclude for zergatik "why" in Basque ist hat it belongs in a different class.

3.7.3. Cues for parameters and the subset trap

What is the most direct evidence for the lexical parameter that distinguishes Basque from Spanish? We note two fundamental and robust ones:

-Separability of wh-words from any quantificational/D value (therefore including f)
-Massive pied-piping

Spanish has only residual Q-wh constructions, such as (83):

(83) Cada cual   se dedica   a   lo suyo
    Each which CL dedicates to his/her business
    'Everyone cares about his/her business'

Not productive at least in Peninsular Spanish (cf. *cada quién, *cada cuando, *cada qué)

Massive pied-piping is a robust phenomenon that shows there cannot be a Q-probe in the lexical configuration of the wh-phrase. The Q-feature is an adjunct. Now, from the point of view of a child that is a sequential or an early bilingual with Basque dominance, and that has set the initial parameter in the Basque way:

(84) [$_{XP}$ Q [$_{XP}$…wh…]]

What is the evidence that can make him/her retreat from this option? The Spanish pied-piping cases are a subclass of the Basque ones, and therefore they are included in the set of potential structures that the Basque setting allows. This is a typical subset trap (Berwick, 1985; Manzini and Wexler, 1987). In the absence of negative evidence, positive evidence of the Spanish type is not enough to reset the initial choice. All Spanish pied-piping cases are compatible with the Basque grammar. The Basque setting therefore is maintained for Spanish in the bilingual mind.

Ancillary assumption: the non-adjacency cases must be relatively infrequent in colloquial Spanish. There is no count on that, but the intuition is that they are infrequent. Also, Paradis and Genesee (1996) « it seems reasonable to conjecture that bilingual children have their input space divided, so their frequency of exposure to each language at any given time is smaller than that of monolinguals acquiring each language ».

The Basque setting applies across-the-board (also in Spanish). An instance of *Input Generalization*.

4. IN-SITU WH-PHRASES IN LABOURDIN BASQUE (DUGUINE AND IRURTZUN, 2014).
4.1. Young Labourdin Basque

Remember the basic word order in Basque, as well as the adjacency requirement imposed on focal elements:

(85) a. Pellok   gereziak jan ditu
        Pello.ERG cherries eaten has
        "Pello ate the cherries"

      b. Nork   jan ditu gereziak?
        Who.ERG eaten has cherries
        "Who ate the cherries?"

c. Zer   jan   du Pellok?
    What eaten has Pello.ERG
    "What did Pello eat?"

The classical analysis of the adjacency requirement (Ortiz de Urbina, 1989; Irurtzun, 2007) has the wh-phrase move to a left peripheral position, say C, followed by movement of the verbal complex to the head of that projection:

(86) [$_{CP}$ nork jan+ditu [$_{IP}$ (jan+ditu) gereziak]]?

Verb fronting following wh-movement accounts for why the embedded clause also shows inversion in successive cyclic movement:

(87) a. Nork   esan du Jonek edan duela   ura?
        Who.ERG said has Jon.ERG drunk has.COMP water
        "Who did Jon say that had drunk water?"

      b. Nork esan du Jonek [$_{CP}$ (nork) edan du-ela [$_{IP}$ ...ura... (edan+du)]]?

A very different pattern has emerged in young speakers of Labourdin Basque. In the speech of those young speakers, along with the standard construction in (85,87), we also find wh-constructions in which the wh-word and the verbal complex are not adjacent:

(89) Nork   gereziak jan ditu?
      who.ERG cherrries eaten has
      "Who ate the cherries?"

In principle a sequence such as (89) has two possible parses: either both the wh-phrase and the verb are in-situ, or the wh-phrase is leftward moved without the verb following. But if we expand the data range we see that the second option is not possible:

(90) a. *Zer Jonek   jan du?
          what Jon.ERG eaten has
          "What did Jon eat?"

      b. ??Zer Pellori   eman dakozu?
          what Pello.DAT given you.have
          "What did you give to Pello?"

      c. Nori   gereziak eman dazkozu?
        who.DAT cherries given you.have
        "Who did you give the cherries to?"

Generalization: the wh-words occupy the unmarked place they would have in an ordinary declarative sentence (S-IO-O-V-Aux).

4.2. French in-situ questions

Duguine and Irurtzun (2014) note that this strategy shares many properties with French in-situ questions, such as (91):

(91) Tu as   vu qui?
       You have seen who
       "Who have you seen?"

Boskovic (1998, 2000) and Mathieu (1999) have noted that wh-in-situ in French displays intervention effects with negation:

(92) a. *Jean ne mange pas quoi?
          Jon NEG eats not what
          "What doesn't Jean eat?"

      b. Qu'est-ce que   Jean ne mange pas?
         what.EST-CE QUE Jean NEG eats not
         "What doesn't Jean eat?"

Likewise, wh-in-situ is ungrammatical in wh-islands, whereas arguments can undergo wh-movement out of them (Mathieu, 1999; Shlonsky, 2013):

(93) a. *Tu te demandes [comment aider qui]?
          You cl wonder   how   to.help who
          "Who do you wonder how to help?"

      b. ?Qui te demandes-tu [comment aider (qui)]?
         Who cl wonder-cl   how   to.help
         "Who do you wonder how to help?"

Young Basque Labourdine displays the same sensitivity to intervention:

(94) a. *Jonek ez du zer jaten?
          Jon.ERG NEG AUX what eat.ASP
          "What doesn't Jon eat?"

      b. Zer ez du Jonek   jaten?
         what NEG AUX Jon.ERG eat.ASP
         "What doesn't Jon eat?"

And to wh-islands:

(95) a. Ez dakizu [nola nori   opari bat eskeini]?
          NEG you.know how who.DAT present one offer
          "Who don't you know how to give a present to?"

      b. Nori   ez dakizu   nola eskeini opari bat?
         who.DAT NEG you.know how to.offer present one?
         "Who don't you know how to give a present to?"

Another characteristic property of wh-in-situ constructions in French is that they can be embedded within strong islands (Obenauer, 1994; Shlonsky, 2013). Compare with extraction:

(96) a. *Qu'est-il   tombé sur la solution en faisant?
          What.is-he fallen on the solution in doing
          *"*What has he fallen on the solution by doing?"*

      b. Il est tombé sur la solution en faisant quoi?
         He is fallen on the solution in doing what
         "He has fallen on the solution by doing what?"

YLB shows the same asymmetry (Duguine and Irurtzun, 2014: 9)

(97) a. *Nori   piztu dute jendearen kexua [_AdvP_ (nori) etxea kentzean]?
          Who.DAT light aux people's anger   house remove.when
          *"*Who did they light people's anger when they took the house to?"*

      b. [_AdvP_ Nori   etxea kentzean]   piztu dute jendearen kexua?
              Who.DAT house remove.when light aux people's anger
         "They lit people's anger when they took the house to whom?"

YLB in-situ questions present the following properties:

(i)   The wh-phrases occupy the same position they would occupy in an unmarked declarative sentence
(ii)   They show intervention effects with negation
(iii)   They show wh-island effects
(iv)   They are possible in strong islands (with pied-piping)

Conclusion: something like the French in-situ strategy has been transferred into Basque (in contact with French). Note that this convergence cannot be established n terms of E-language correspondences. What is "transferred" is a rule: don't move your wh-s. Superficially, the order that results from applying this rule to Basque maximally separates Basque from French in terms of word order.

(98) a. Qu'as t-il acheté?   (Movement into C)
      b. Zer erosi du Jonek?   (Movement into C)
         what bought AUX Jon.ERG
         "What did he/she buy?"

(99) a. Jonek   zer erosi du?   (in-situ)
         Jon.ERG what bought has
         "What did Jon buy?"

      b. Jean a   acheté quoi?   (in-situ)
         Jean has bought what
         "What did Jean buy?"

How may this transfer have come about? Why is it that unlike the Basque-Spanish cases, the Basque setting for the parameter (move wh-phrases) is not enough to discard a French parse of the data in terms of a wh-in-situ representation? This is a genuine innovation in the grammar of Basque, but one that leaves the previous system alive. YBL is clearly a superset of the previous Basque grammar (see Biberauer and Roberts, 2009): A Move AND in-situ language.

4.3. Ambiguous input in Basque partial questions

The starting point of Duguine and Irurtzun's proposal: "one of the crucial factors triggering the change is the abundance of critically underspecified data in the PLD that Labourdin Basque learners have to parse" (p.12). They mention three properties of Basque that give rise to such an ambiguity: (i) the generalized pro-drop system of the language, coupled with its SOV status; (ii) ambiguous parse of intransitve sentences; (iii) topicalization patterns.

For the first property, consider the fact that all arguments in Basque can be null:

(100) a. Zuk   Mireni   gereziak   eman dizkiozu
         you.ERG Miren.DAT cherries.ABS given you.have.to.her
         "You gave (the) cherries to Miren"

      b. *pro pro pro* eman dizkiozu
                       given you.have.to.her
         "You gave them to her"

Generalized pro-drop does not correlate with finiteness in Basque. It is also available in clausal nominalizations, which like finite forms:

(101) [Clausal NomP *pro pro pro* ematea]   keinu polita da
                                 give.NOM.DET gesture nice   is
         "That you gave it to her is a nice gesture"

Now consider a sentence such as (102):

(102) Zer eman duzu   zuk?
      what given you.have you.ERG
      "What did you give?"

The basic cue telling us that both the wh-word and the verb have moved to C is the overt ergative subject, left behind by these movements. But if we eliminate the overt subject from the sentence (pro), then this cue disappears, and the string can be parsed as one where the arguments are in-situ in an basic SOV clausal structure:

(103) [[[[...*zer*...vP] *eman* AspP] *duzu* TP] Cwh] ?

The ambiguity of the input extends to intransitive clauses. The neutral order of Basque for intransitives is SV. This order will remain unchanged whether the subject is a wh-phrase or not, under any analysis (same for unergatives):

(104) a. Nor hil da?   b. Jon hil da
         Who died is   Jon died is
         "Who died?"   *"Jon died"*

A third property of Basque that generates structural ambiguities in wh-questions is the syntax of topicalization. The topic position stands above the wh-position in Basque:

(105) a. Gereziak nork   jan ditu?
         Cherries who.ERG eaten has
         "Who ate (the) cherries?"

      b. *Ezer,   nork   jan du?
         Anything, who.ERG eaten has

In certain configurations, the availability of topicalization will trigger ambiguity. Take (84):

(106) Mirenek   zer jan du?
       Miren.ERG what eaten has
       "What did Miren eat?"

(106) is compatible with an analysis in which both the wh-word and the verbal complex have moved to C, but it is also compatible with an in-situ position for the wh, and an IP internal position for the subject:

(107)   [[*Jonek* [[...*zer*...$_{vP}$] *jan* $_{AspP}$] *du* $_{TP}$] C$_{wh}$] ?

Duguine and Irurtzun (2014) quantify the amount of ambiguous input in a CHILDES corpus (the Luque corpus) composed of informal and spontaneous (non-elicited) conversations (23.833 utterances in total) between teachers and caretakers and 38 children aged from 2 to 4. This is not a labourdin corpus. It helps to set the base of ambiguous input *before* the development of the in-situ strategy. The corpus presents unambiguous examples such as (86a,b), and ambiguous strings such as those in (108a,b).

Unambiguous input

(108) a. Nork   egin dau puzzlea?
          Who.ERG done has the.puzzle
       "Who made the puzzle?"

       b. *A ver*,   zelan abesten dau igelak?
          Let's see, how sing.ASP aux frog.ERG
       "Let's see, how does the frog sing?"

Ambiguous input

(109) a. Burua aterata   nork   deko?
          Head stick.out who.ERG has?
       "Who has the head stuck out?"

       b. Hau   norena da?
          this.ABS whose is
       "Whose is this?"

Figure 1: Primary linguistic data of wh-questions in Basque (western varieties)
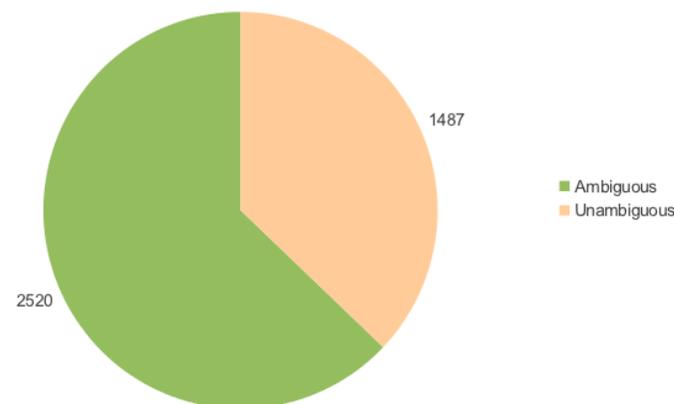
This is suggestive, but as Duguine and Irurtzun point out, it is not enough to trigger grammatical change. If this were enough, all Basque dialects would have developed in-situ wh-questions. The extra factor that intervenes here is contact with French.

4.4. The sociolinguistic factor

Basque is a "severely endangered" language (UNESCO) in the French (Northern) Basque Country. According to the Basque Government survey of 2013, slightly more than a fifth of the population speaks Basque in the Northern Basque Country.

-21,4% bilinguals
-9.1% passive bilinguals
-69,5% French monolinguals

Figure 2: L1 by age in 2011 (from Duguine and Irurtzun, 2014)

L1 by Age – 2011 (%)

		variable
		Basque & French
		Basque

>=65   50–64   35–49   25–34   16–24

30.2
23.2
15.5
9.9
7.7
5.6
4.2
5.1
7.5
5.7
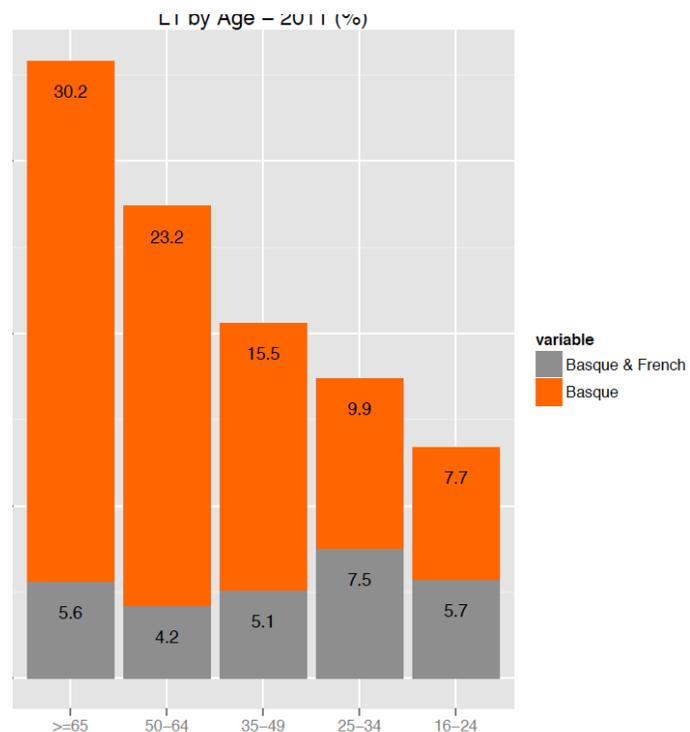
Figure 3: Typology of Basque speakers in terms of fluency (same source)

Fluency of Bilingual Speakers – 2011 (%)

		variable
		Basque–prominent
		Balanced
		French–prominent

>=65   50–64   35–49   25–34   16–24

26.7   35.1   37.9   40.2   49.6
41.8   42.2   46.9   42.7   38.5
31.5   22.7   15.2   17.1   11.9
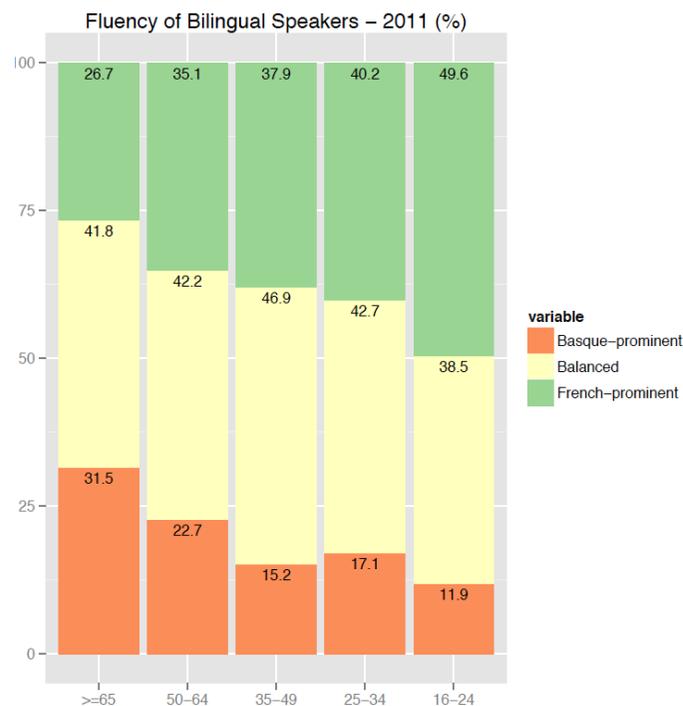
4.4. The in-situ setting

Duguine and Irurtzun propose that there is a crucial factor in the transfer of the in-situ strategy to Basque. They suggest that the in-situ strategy is computationally simpler than the one involving displacement. They point at work by Jakubowicz (2011) according to which children acquiring French (both typically developing children and children with SLI) attempt to avoid those syntactic patterns that involve a long-distance relation between the surface and the thematic position of the wh-phrase by resorting to in-situ constructions that can be target-deviant. English speaking L2 learners of French show preference for in-situ wh-questions over movement-based wh-questions, even though the in-situ strategy is not available in their language (this is questionable) (Scheidnes and Tuller, 2010). Creoles tend to choose in-situ questions if the mixed input to creolization presents this strategy (Clements and Mahboob, 2000). Again, this cannot be the trigger of the change.

18

Could transfer follow from Feature Economy?

Consider (21) again, repeated here:

(21)   Given an adequate structural representation R for a substring of input text of the PLD S, keep postulating R for new incoming input of a language S'. Add R' only if forced.

Assume R corresponds to moving the wh-phrase, an option available in French. Upon hitting the Basque unambiguous examples, the bilingual learner can keep using R as a means to parse the input from S'. The setting R is thus retained. The logic of the Basque influence in wh-word-verb adjacency in Spanish follows this logic of retention.

Consider now the alternative setting R', corresponding to the in-situ strategy, available in French too. The bilingual learner can use R' to parse a substantial amount of S' input, the one that corresponds to ambiguous input. This result is adding R' as an additional possible setting for S'.

Handout II. Contact Issues in Case and Agreement. Differential Object Marking in Basque
Ricardo Etxepare (CNRS IKER UMR 5478)

1. DOM in Spanish

In contact with Spanish, a Differential Object Marking (DOM) language, dialects of Basque in contact with Spanish have also developed DOM (Rezac and Fernandez, 2016 ; Rodriguez-Ordoñez, 2016 ; Odria, 2017, among others; Austin, 2014). DOM describes the fact that under certain semantic conditions having to do with the referential status of the object, this object arises with an overt case marking. In Spanish, this marking is the dative preposition *a*. The relevant deictic parameters that allow DOM in Spanish are animacy and specificity (data from the variety of Spanish spoken in the Basque Country, a *leísta* dialect) (from Romero and Ormazabal, 2013):

(1)   a. Te   he   visto a   ti
          2P.DAT.CL I.have seen PREP you
          'I have seen you'

        b. Le   he   visto *(a) la niña
          3P.DAT.CL I.have seen   PREP the girl
          'I have seen the girl'

        c. Le   enviaron (a)   un especialista a   su casa
          3P.DAT.CL sent   PREP a specialist   PREP his house
          'They sent a specialist to his house'

(2)   a. Busco   un médico
          I.look.for a doctor
          'I am looking for a doctor'

        b. Busco a   un médico
          I.look PREP a doctor
          'I am looking for a (particular) doctor'

(3)   a. He   visto uno
          I.have seen one
          'I have seen one (object/non-specific human)'

        b. Le   he   visto a   uno
          3P.DAT.CL I.have seen PREP one
          'I have seen one (human specific/*object)'

DOM extends to other types of objects whose status as specific is not clear, and for which animacy looks like the only option:

(4)   a. No le   he   visto *(a) nadie
          NEG 3P.DAT.CL I.have seen   PREP anyone
          'I haven't seen anyone'

        b. No se   ha   visto *(a) sí mismo en el espejo
          NEG 3P.CL I.have seen PREP SE self   in the mirror
          'He/she didn't see himself/herself in the mirror'

(5)   DOM: add a marked case (dative) to objects showing either of the following high features in the referentiality scale: [animate] AND/OR [specific].

2. DOM in Basque

Some Basque dialects, and only those that are in contact with Spanish, have developed Differential Object Marking in their objects (Fernandez and Rezac, 2016 ; Rodriguez, 2016 ; Odria, 2017, among many others). Thus ordinary transitive predicates in Basque present an ERGATIVE-ABSOLUTIVE alignment, with the transitive subject being marked Ergative and the object being (morphologically) unmarked (absolutive)

(6)   a. Nik zu   ikusi zintudan   (Standard)
          I.ERG you.ABS seen I.have.you.PAST
          'I saw you'

        b. Nik zuri   ikusi nizun   (Dialectal)
          I.ERG you.DAT seen I.have.it.to.you.PAST
          'I saw you'

Whereas the object of transitives in standard Basque is unmarked (absolutive), it can show up with dative case marking in the relevant dialects. This dative case also triggers dative agreement in the auxiliary (which corresponds to the ditransitive one). The structure is parallel to the DOM structures existing in the Spanish spoken in the Basque Country, with the clitic *le* the equivalent of dative agreement in Basque.

Despite this highly transparent structural equivalence, Basque does not allow DOM in all those cases in which it is possible in Spanish. For instance, in many varieties, DOM is only possible with 1st and 2nd person (in Dima, Arratia and Ulzama, see Zuazo, 2003; Mounole, 2011; Fernandez and Rezac, 2016). Something like the Spanish equivalent of (1b) would be impossible in those varieties :

(7)   *Neskari ikusi diot
          girl.DAT seen PRES.ROOT.3DAT.1ERG
          'I have seen the girl'

Even in those varieties in which (7) is not completely discarded, it is less common or less preferred than ½ person. Specificity, on the other hand, does not play a role in Basque DOM. So (8) is not possible either:

(8)   *Aditu   bati   bialdu zioten
          specialist one.DAT sent   AUX
          'They sent a specialist' (cf. Spanish *Enviaron a un especialista*)

Also, Basque DOM is impossible with (4a,b) in those same varieties. Both Odria (2017) and Rodriguez (2017) have found equivalents of (7) in spontaneous speech in Gernika, although the same speakers tend to declare those ungrammatical. The following table synthesizes the distribution of DOM in Basque.

(9)   Comparing DOM in Basque and Spanish

	Spanish	Basque
1/2 person	OK	OK
Animate Definite	OK	%OK
Animate Specific	OK	*
Animate non-specific	OK	*
Non-animate specific	OK	*
Animate non referential	OK	*

Here's the question: The direct evidence that informs the rule of DOM for a (bilingual) Basque speaker is based on the occurrences of dative marked DPs in Spanish. What prevents the Basque learner to generalize DOM to all those instances in which it is possible in Spanish, based on his/her acquired knowledge of Spanish?

The existence of variable referential conditions in the distribution of DOM cross-linguistically is obviously a well known fact. Languages seem to restrict DOM to varying subsets of the feature combinations expressed in (9) (see Dalrymple and Nikolaeva, 2011). Examples :

Russian (from Bossong, 1991:160): only the animacy of the object matters

(10)   a. Ja vstreëaju   dorogix   gostej
          I receive.1SG dear.GEN.PL guest.GEN.PL
          'I receive beloved guests'

b. Ja pokupaju dorogije   vešči
    I buy.1SG dear.NOM.PL thing.NOM.PL
    'I receive expensive things'

Turkish (Enç, 1991: 4-5): only specificity. The accusative marking *–(y)i* correlates with the specific interpretation of the object.

(11)   a. Ali bir pianoyu   kiramalak istiyor
          Ali one piano.ACC rent.INF   wants
          'Ali wants to rent a certain piano'

           b. Ali bir piyano kiramalak istiyor
             Ali one piano rent.INF   wants
             'Alia wants to rent a piano (anyone would do)'

Hebrew (Danon, 2006: 3): only definiteness.

(12)   a. Dan kara *(et) ha-itonim
          Dan read DOM the.newspapers
          'Dan read the newspapers'

           b. Dan kara (*et) (kama) itonim
             Dan read DOM some newspapers
             'Dan read (some) newspapers'

Hindi (Mohannan, 1994: 80, 85): both animacy and specificity trigger the presence of DOM. *-ko* identical to the dative marker.

(13)   a. Ilaane bacce-ko uTaayaa
          Ilaa.ERG child-DOM lift.PERF
          'Ila lifted a/the child'

           b. Ilaane haar   uTaayaa
            Ilaa.ERG necklace lift.PERF
           'Ila lifted a/the necklace'

From this point of view, there is nothing particularly noteworthy in the Basque cases, beyond what is due to account for this variation cross-linguistically. But the contact situation has the potential to underline specific inductive problems that L1 acquisition doesn't have to adress. Why is it that the effects of contact stop at that precise point ? Why isn't it the case that contact-induced change goes all the way in Basque to mimic the

Spanish system ? Let's call this the **failed generalization problem** of contact-induced change.

3. Syntactic Conditions on DOM

The restriction in Basque DOM as it compares to Spanish DOM, the source of the grammatical change in Basque, has an arbitrary character from the point of view of E-language. But what if the correspondence between the grammars of Basque and Spanish are not established on the basis of surface similarities or constructional parallelisms ?

3.1. Raising to Object

Romero and Ormazabal (2013) : DOM as a condition on agreement/case. ECM:

(14)   Vimos **al**   barco estrellarse contra los arrecifes
         we.saw PREP.DET ship dash.CL   against the reef.PL
         'We saw the ship crash on the reef'

(15)   a. Vimos **al** avión estrellarse contra la montaña
            saw.1pl A-the plane crash.down against the mountain
         'We saw the plane crash into the mountain'

         b. *Vimos **el** avión estrellarse contra la montaña
            saw.1pl the plane crash against the mountain
         'We saw the plane crash into the mountain'

         c. Vimos estrellarse **el** avión contra la montaña
            saw.1pl crash the plane against the mountain
         'We saw the plane crash into the mountain'

In Northern Peninsular leísta dialects, where 3rd person animates are doubled by the agreement clitic *le* (see section 2 above and references), Raising-to-Object DOM is accompanied by doubling with the agreement-clitic, minimally contrasting with the impossibility of clitic doubling when the DP is not marked with DOM, as expected (see also Zdrojevsky 2008; Ormazabal and Romero 2013a for Rioplatense Spanish).

(16)   a. **Le** vimos **al** avión estrellarse contra la montaña
            3s saw.1pl A-the plane crash against the mountain
         'We saw the plane crash into the mountain'

         b. (*Le) vimos estrellarse **el** avión contra la montaña
            3s saw.1pl crash the plane against the mountain
         'We saw the plane crash into the mountain'

Interestingly, Clitic Doubling in Northern Peninsular varieties (the Spanish oft he Basque Country) is only possible in Raising to Object. It is not otherwise possible with the bare transitive structures in (17):

(17)   *Le   has   visto al   barco?
          3S.DAT you.have seen PREP.DET ship
         "Have you seen the ship"

Laca (1995) notes that similar non-animate DOM objects arise in causative constructions.

(18)   El mago   hizo levitar   **a**   las sillas
         the magician made levitate.INF PREP the chairs
         "The magician made levitate the chairs"

And Small Clause predications (Laca, 1985; apud Ormazabal and Romero, 2013):

(19)   a. Luis dejó **a**   ese árbol sin una aceituna
            Luis left PREP that tree without an olive
         'Luis left that tree without a single olive'

         b. Veía   distante **a**   la ciudad
            I.watched far.away PREP the city
         "I was watching the city in the distance"

Compare *Veía a la ciudad*

If Ormazabal and Romero are right, DOM reflects a Case/Agree relation between v and the DOM object.

(20) $[v_{(iT)}...V...DP\text{-}DOM_{(uT)}]$

In Basque, equivalents of (14), (18) and (19) are not possible with DOM:

(21)   a. *Itsasontziari uretan   hondoratzen ikusi genion
            ship.DET.DAT waters.LOC sink.ASP   seen 1.ERG.ROOT.3SA.3SDAT.PAST
         "We saw the ship sink in the waters"

         b. *Aztiak   aulkiei   lebitatu-arazi zien
            magician.ERG chair.DET.DAT levitate-cause PAST.3S.ERG.ROOT.3SA.3PL.DAT
          "The magician caused the chairs to levitate"

c. *Hiriari   urrutian   ikusten zion
   city.DAT distant.LOC see.ASP PAST.3S.ERG.ROOT.3SA.3SDAT
   "He/she watched the city in the distance"

The way you would say that is by using ordinary transitive structures of the ergative-absolute sort:

(22)   a. Itsasontzia   uretan   hondoratzen ikusi genuen
        ship.DET.ABS waters.LOC sink.ASP   seen 1.ERG.ROOT.3SA.PAST
        "We saw the ship sink in the waters"

        b. Aztiak   aulkiak   lebitatu-arazi zituen
           magician.ERG chair.DET.ABS levitate-cause PAST.3S.ERG.ROOT.3PL.ABS
            "The magician caused the chairs to levitate"

        c. Hiria   urrutian   ikusten zuen
           city.ABS distant.LOC see.ASP PAST.3S.ERG.ROOT.3SA.3SDAT
           "He/she watched the city in the distance"

But alternations similar to Spanish in the nature of Case marking can be found in Basque, even in those dialects which do not present DOM. For instance, the equivalents of the arbitrary SE constructions of Spanish, like (23a), have a natural dative object counterpart in Basque:

(23)   a. Se le   vio a   Javier leer   un libro una vez
           SE 3S.DAT saw PREP Javier read.INF a book one time
           "Javier was seen to read a book once"

        b. Xabierri   liburu bat   leitzen ikusi zitzai**o**n   aspaldian
           Xabier.DAT book one.ABS read.ASP seen AUX.3S.ABS.**3S.DAT** long.time.ago
           "Xabier was seen to read a book long time ago"

(23) contains an animate object, which could be the target of DOM independently of the structure for at least some speakers. But the important thing is that dative case marking arises also in those varieties in which DOM is not operative. Basically the equivalents of arbitrary SE structures in Spanish have raising of the object (to Subject?), which is marked by DOM:

(24)   Xabierri   ez   zaio   aspaldian ikusten
        Xabier.ERG NEG ROOT.3S.ABS.3S.DAT long time see.ASP
        "Xabier hasn't been seen for a long time"

Those structures require dative clitic doubling in Spanish:

(25)   A   Javier no se *(le)   ve desde hace tiempo
        PREP Javier NEG SE 3S.DAT see since long time
        "Javier hasn't been seen for a long time"

The structure can be reproduced with non-animate objects in Raising to Object contexts. In this case, there is no DOM outside the raising cases:

(26)   a. Itsasontziari uretan hondoratzen ikusi zitzaion
           ship.dat waters.loc sink.asp seen AUX.3S.ABS.3S.DAT
           "The ship was seen to sink in the waters"

        b. *Itsasontziari ikusi zitzaion
           ship.DAT   seen AUX.3S.ABS.3S.DAT
           "The ship was seen"

That arbitrary subjects of the SE-sort are crucial here is shown by the following contrast, in which the case-status of the object depends on the aspect of the clause.

(27)   a. *Hegazkina   uretan   hondoratzen ikusi zen
           Plane.DET.ABS water.in sink.ASP   seen was
            "The plane was seen to sink in the waters"

        b. Hegazkina   uretan   hondoratzen ikusten da   (argazki horretan)
           plane.DET.ABS waters.LOC sink.ASP   see.ASP is picture that.loc
           "The plane was seen to sink in the waters in that picture"

Where imperfective aspect seems to license an arbitrary/generic subject, which otherwise needs the support of a raising object (which then requires dative case marking). (27b), unlike (26a), does not require clitic doubling:

(28)   a. En esa fotografía, se ve   al   barco hundirse
           in that picture,   CL sees PREP.DET ship sink.SE
           "In that picture, the ship is seen to sink"   (imperfective)

        b. En esa fotografía, se *(le)   vio hundirse al   barco
           in that picture,   SE 3S.DAT saw sink.SE PREP.DET ship
           "In that picture the ship was seen to sink"   (perfective)

Summary

In both DOM and non-DOM varieties of Basque there is Differential Object Marking in at least a subset of the Raising to Object/Subject? Constructions. Those cases seem to

correlate with the presence of a dative clitic in their Spanish counterparts. Obligatorily, at least in Basque Spanish.

3.2. DOM restrictions and the PCC

If we focus on the possible and impossible cases along the "referential scale", we will find that the objects that admit DOM are the same that trigger PCC (Person-Case) effects in ditransitive clauses.

PCC (Bonet, 1991:36)

(29)   Person-Case Constraint: if DAT, then ACC/ABS=3rd person

In the Romance domain, the PCC is also known as the *Me-Lui* constraint (Perlmutter, 1971), a restriction that operates on clitic clusters. Compare (30a,b):

(30)   a. Pedro te   lo   envía
           Pedro 2DAT 3ACC he.sends
           "Pedro sends it to you"

         b. *Pedro te   me   envía
            Pedro 2DAT 1ACC he.sends
           "Pedro sends me to you"

The PCC also restricts agreement systems (Albizu 1997). Basque (31a,b):

(31)   a. Zuk   niri   liburua   saldu d-Ø-i-da-zu
           you.ERG me.DAT book.the.ABS sold   PRES-3ABS-ROOT-1SDAT-2S.ERG
           "You sold me the book"

         b. *Zuk   etsaiei   ni   saldu naiezu
            you.ERG enemies.DAT me.ABS sold   1ABS-ROOT-3PL.DAT-2S.ERG
           "You sold me to the enemies"

Is this a morphological or a syntactic restriction? It is worth noting that in Basque, the PCC seems not to arise in non-finite contexts:

(32)   [Zuk   ni   etsaiari   saltzea]   eskandalagarria da
        you.ERG me.ABS enemy.DAT sell.nom.det scandalous   is
        "It is scandalous that you should sell me to the enemies"

This type of asymmetry has been adduced as an argument in favour of a morohological approach to the PCC (Albizu, 1997; Bonet, 1991). But in fact, you may notice that the order of the constituents is not the unmarked one in (32). If we invert the relative order of absolutive and dative, the sentence is degraded for me also in non-finite sentences:

(33)   ??[Zuk   etsaiari   ni   saltzea]   eskandalagarria da
          you.erg enemy.dat me.abs sell.nom.det scandalous   is
          "That you should send me to the enemies is scandalous"

The order in (33) is the one that corresponds to an ordinary declarative sentence in Basque. The other order, the one in (32) is the one that corresponds to the repair order in cases of PCC conflict. Speakers of Basque resort to several strategies to cope with a PCC configuration. One of them is not to agree with the dative, in which case the dative argument follows the absolutive one:

(34)   Zuk   ni   etsaiari   saldu nauzu
        you.ERG me.ABS enemy.DAT sold 1S.ABS.2S.ERG
        "You sold me to the enemy"

So I will assume that the PCC is related to Case/Agree.

As Ormazabal and Romero (2007) note, the actual restriction does not involve person, but animacy. Consider (35):

(35)   a. Pedro te   lo   envía
           Pedro 2DAT 3ACC he.sends
           "Pedro send it to you"

         b. *Pedro te   le   envía
            Pedro 2DAT 3ACC(ANIMATE) he.sends
           "Pedro sends him/her to you"

That the constraint has nothing to do with clitic clusters in and of itself is shown by the contrast in (36)-(37) (from Ormazabal and Romero, 2007:15-16)

(36)   a. Mateo$_i$ piensa que lo$_i$ entregaste a la policia
           Mateo thinks that 3ACC handed to the police
           "Mateo thinks that you handed him to the police"

         b. *Mateo$_i$ piensa que se   lo$_i$ entregaste a la policia
            Mateo   thinks that 3DAT 3ACC(ANIMATE) you.handed to the police
           "Mateo thinks that you handed him to the police"

(37)   a. El paquete_i especifica que lo_i   entregues al portero
        the package specifies   that 3ACC hand.you to.the doorman
        "The package specifies that you should hand it tot he doorman"

        b. El paquete_i especifica que se   lo_i   entregues   as portero
        the package specifies that 3DAT 3ACC(INANIMATE) you.handed to the doorman

In Basque too, the restriction is a bit more complicated in the sense that 3rd person may not always rescue the structure. Third person overt pronouns behave as ½ person:

(38)   a. *Zuk   Mikeli   ni   aurkeztu naiozu
        you.ERG Mikel.DAT me.ABS sold   1ABS-ROOT-3PL.DAT-2S.ERG
        "You introduced me to Mikel"

        b. ??Zuk   Mikeli   bera   aurkeztu diozu
        you.erg Mikel.dat him.abs introduced PRES.ROOT.3DAT.1ERG
        "You introduced HIM to Mikel"

Ormazabal and Romero propose to formulate the relevant generalization as in (39):

(39)   a. Person-Case Constraint: if DAT, then ACC/ABS=[-animate]
        b. If object agreement encodes animacy, no other argument can be licensed through verbal agreement (including clitics in verbal agreement)

Ormazabal and Romero derive (39) from two independent principles that affect v-relations:

(40)   a. Object Animacy Generalization: object relations, in contrast to subject and applied object relation, are sensitive to animacy (that is, if you have an animate object, and v is a probe, it will be sensitive to animacy)

        b. Object Agreement Constraint: If the verbal complex encodes object agreement, no other argument can be licensed through verbal agreement.

The verbal complex may only encode one agreement relation. In ditransitive constructions where more than one potentially agreeing element appears, only one can in fact agree with v. This is at the basis of the PCC. The Basque cases show in any case, that animacy is not enough.

There is a perfect correlation between the type of DP that participates in PCC constraints in Basque and the DPs that can be marked by DOM in Basque. Besides 1 and 2 person, there are also proper names and overt pronouns, in many of those varieties:

(41)   Berari   ikusi diot
        him.DAT seen **I.have.it.to-him**
        "I saw him"

If Romero and Ormazabal are right, the DOM marked DPs must enter an Agree relation with v. One thing that we could say is that the relevant feature here is animacy as it is expressed in D (pronouns). There is an issue as to the status of null pronouns, which can be the object in *introduce* type predicates:

(42)   A: Xabier etorri da
          Xabier come is

        B: Zuzendariari aurkeztu   diozu?
          Director.DAT   introduced PRES.ROOT.3DAT.1ERG
        "Did you introduce him to the director?"

3.3. Person behaves like a clitic in Basque

There is a clear divide in Basque between Person and Number agreement (Etxepare, 2003, 2006, 2012, among others). Person agreement behaves as a clitic. Number agreement doesn't. This can be seen in Long Distance Agreement:

(43)   a. [Presidenteari   _ aurkeztea]   pentsatu zaituzte
          president.DET.DAT   introduce.NOM.DET planned AUX.2ABS.3PLE
        'They have planned to introduce you to the president'

        b. *[Presidenteari   _ aurkeztea]   pentsatu dituzte
          president.DET.DAT   introduce.NOM.DET planned AUX.2ABS.PL.3PLE
        'They have planned to introduce them to the president'

Long Distance Person agreement across a dative argument is possible. Long Distance Number Agreement is not. Without the dative argument, the sentence is good:

(44)   [_ aurkeztea]   pentsatu dituzte
          introduce.nom.det planned aux.2A.3plE
        'They have planned to introduce them'

In the abstract structure (43), with the dative DP as an intervener person can circumvent the intervention condition by clitic climbing into the higher T, via the outer edge of v or internal T, and therefore allowing the number probe in the higher v to agree with the absolutive DP inside the dependent clause. Number agreement surfaces in the higher finite verb. This possibility is not available to DPs without a person feature.

(45)   ...T...(**DP**_**DAT**) v [... v...(**DP**_**DAT**) v..(**DP**_**DAT**)...]

Number/Person generalization in Basque (Etxepare, 2012):

(46)   a. Person affixes in the finite verb are clitics in Basque
            b. Number affixes are agreeing probes

3.4. DOM objects behave as DPs (unlike other datives)

As in many languages, and for reasons we don't understand very well, indirect objects fail to license secondary predicates:

(47)   Mirenek_i   Peruri_j   liburuak   pozik_{i/*j} laga dizkio
            Miren.ERG Peru.DAT books.DET.ABS happy   lend she.aux.it.to.him
            "Miren_i lend the books to Peru_j happily_{i/*j}"

The same phenomenon occurs in Spanish (Demonte, 1995), and it has been related to the prepositional status of the controller (I don't believe a word):

(48)   a. María le   ha dejado a   Pedro los libros contenta
               María 3s.dat has left   PREP Pedro the books happy
            "Maria_i left the books to Pedro happy_i

            b. *Maria le ha dejado a Pedro los libros contento

            c. *Maria ha repartido regalos entre   sus amigos contentos
               María has distributed presents among her friends happy.PL
            "Maria has distributed presents among her friends happy"

Odria (2017), and Fernandez and Rezac (2016) observe that DOM marked objects behave as absolutive object DPs in licensing secondary predication (Larrabetzu Basque, Biscay). Data from Odria (2017):

(49)   a. Zuk_i   ni_j   ostendute_{i/j} harrape nozu
               you.ERG me.ABS hidden   catched you.aux.me
            "You caught me hidden"

            b. Zuk_i   niri   ostendute harrape dostezu
               you.ERG me.DAT hidden   caught you.aux.it.to.me
            "You caught me hidden"

Elgoibar Basque (Biscayan variety):

(49)   Geldi_i ikusten diazu   niri_i?
            Still   see.ASP you.aux.it.to.me me.DAT
            "Do you see me still/motionless (by a chance)?"

Interestingly, DOM marked objects align in this regard with so-called "high datives":

Dative subjects of causatives

(50)   Amak_i   haurrari_j gaixorik_i joan-arazi zion   eskolara
            Mother.ERG child.DAT sick   go.cause she.aux.it.to.him school.to
            "The mother_i made the child go sick_i to the school"

Possessor raising

(51)   Gurasoek_i umeari_j   txupetea_j   negarrez_{i/j} kendu   diote
            Parents.ERG child.DET.DAT pacifier.DET.ABS crying   removed they.aux.it.to.him
            "The parents removed the pacifier from the child crying"

Experiencer datives

(52)   Mireni_i   mutil hori_j   mozkortuta bakarrik_i gustatzen zaio
            Miren.DAT boy   that.ABS drunken only like.ASP   he.aux.to.her
            "Miren likes that boy with a shaven head"

"Raising to object"

(53)   Hegazkinari_i uretan osorik_i hondoratzen ikusi zitzaion
             plane.det.DAT water.in of.a.piece sink.ASP   seen AUX.3P.S.ABS-**3P.S.DAT**
            "The plane_i was seen to sink of-a-piece_i in the waters"

What does being a "high dative" mean? It basically means that it is gonna be realized in a position higher than the absolutive. This is not the case with a subset of datives, typically Goal denoting ones. For instance, the unmarked word order for a predicate like *hurbildu* "approach" is one where the absolutive precedes the absolutive:

(54)   a. Jon   Mireni   hurbildu   zaio
               Jon.ABS Miren.DAT approached is.to.her
            "Jon approached Miren"

            b. Mireni   Jon   hurbildu   zaio
               Miren.DAT Jon.ABS approached is.to.her
            "Miren was approached by JON"

In all the cases in (50)-(53), the unmarked order is one in which the dative precedes the absolutive, if we add an object:

(55) a. Amak   haurrari goxokiak utzi-arazi dizkio
      Mother.ERG child.DAT sweets.ABS leave-cause PRESENT.ROOT.3PL.ABS.3S.DAT
      "Mother forced the child to leave the sweets"

   b. Amak   goxokiak haurrari utzi-arazi dizkio
      Mother.ERG sweets.ABS child.DAT leave-cause PRESENT.ROOT.3PL.ABS.3S.DAT
      "Mother forced the CHILD (not someone else) to leave the sweets"

Reflexive licensing:

(56) a. Jon   bere buruari mintzatu zaio
      Jon.ABS his head.DAT talked   is.it.to.him
      "Jon talked to himself"

   b. *Bere buruari Jon   gustatzen zaio
      his head.DAT Jon.ABS like.ASP   is.it.to.him
      "John is pleasant to himself"

   c. Joni   bere burua   gustatzen zaio
      Jon.DAT his   head.ABS like.ASP   is.it.to.him
      "Jon likes himself"

3.5. The analysis

The analysis of DOM must take into account:

   (i)   Raising to Object –Subject?-(with arbitrary subjects)
   (ii)   PCC-triggering objects
   (iii)   The relevant feature (D/1/2 person) belongs in the clitic system
   (iv)   The DOM marked DP aligns with high datives in licensing secondary predicates
   (v)   It is dative-marked

The first three properties suggest that the object enters into an Agree relation with v (T?). The fourth one that it ends up in a relatively high position of the verbal predicate. The last one reminds us that case-wise, this is not unlike an ordinary Goal dative in a ditransitive construction. Goal datives look prepositional. Remember that they can be used to repair a PCC violation by just not agreeing. Let me propose the following underlying representation for DOM (animacy in v optional, as required by Ormazabal and Romero, and as shown by the optionality of DOM):

(57) a. $v_{[animacy; Case]}$...V $[_{PP}$ Ø $[_{CIP}$ [DP K$]_i$ Cl$_i$]]] -> incorporation of P (i=animacy)
      b. ...V+ Ø $[_{CIP}$ [DP K] Cl]]] -> Raising of the clitic to small v, and Merger
      c. ...CL+$v_{[u:animacy; iT]}$...V+Ø $[_{CIP}$ [DP K$]_i$ Cl$_i$]]] -> Raising of KP to outer edge
      d. ...$[_{vP}$ [DP K] CL+$v_{[u:animacy; iT]}$...V+Ø $[_{PP}$ Ø... (High dative behaviour)

Clitic raising (a D category, realized by the dative affix, the only 3rd person affix in the Basque verbal paradigm) is crucial in the process. This excludes all those cases of animacy which are not expressed in D. The P that we see there is the same P that licenses Goals, I will assume. Unlike in Goals, this P incorporates into V, extending the phase to the verbal domain.

The incorporation of the preposition should perhaps be related to another general feature of Basque: animate DPs cannot directly express Grounds in spatial relations:

(58) a. In me   b. In John   c. To me   d. To John

(59) a. *Ni-n   b. *Jon-en   c. *Ni-ra   d. *Jonera
      I-in   Jon-in   I-to   Jon-to

The only way to have an animate DP in a spatial relation is by adding another locative particle, -*ga-,* specific for those occasions:

(60) a. Ni-ga-n   b. Zu-ga-n-a
      I-suffix-in   you-suffix-in-to
      "In me"   "To you" (spatial)

With animate DPs, preposition incorporation would be a type of repair from the impossible configuration in (59a-d). Also, primary spatial adpositions in Basque present D-dropping:

(61) a. Etxe-ra   b. Etxe-an   c. Etxe-tik
      house-to   house-in   house-from
      "To the house"   "In the house"   "From the house"

(62) P is incompatible with the overt realization of D in Basque

Unless you insert more spatial structure, like spatial nouns (Axial Parts, Jackendoff/Svenonius)

(63) Etxearen   aurre-tik
      house.det.gen front.from
      "From the front oft he house"

For Goal datives, I will assume that the verbal semantics is able to provide some extra spatial ground inside the PP, so that animate DPs are good without preposition incorporation. In those cases, only the clitic raises, via the edge of the PP (yielding dative agreement):

So what is "borrowed" from Spanish?

Franco (1995): the distribution of clitic *le* in leista Basque Spanish, always accompanies DOM.

(64)   a. (le)   hemos visto a   un niño   (Standard Spanish)
            3S.DAT we.have seen PREP a child
        "We have seen a child"

        b. *(le)   hemos visto a   un niño   (Basque Spanish)
            3S.DAT we.have seen PREP a   child
        "We have seen a child"

Let us say that the borrowing can be interpreted in the following terms:

(65)   Insert a clitic in the object position (clitic doubling)

The rest follows. Once you add a clitic you must worry about case licensing both the clitic and the DP, so you need a P for the internal DP, and you must raise the clitic outside for Case/Agree. This much is necessary for any dative (indirect object argument). What is special is the animacy restriction related to the DOM. What I would lie to suggest ist hat this narrowing effetc in the intepretation of DOM objects is a result of competition with ordinary objects, which are also possible in those dialects. That is, objects in general can be licensed in the absolutive configuration. If you license an object in some other configuration, let it be to get some special interpretation. In Romero and Ormazabal's approach, that means that you will add a feature [animate] in v. Obligatorily in those cases, and only in those cases.

As for (65), this is I think a clear case of *Input Generalization*.

(66)   Maximize already postulated features

Coupled with the general principle in (67), which says that you will treat the two languages alike:

(67)   Avoid constructing more than one grammar

Clitic doubling being possible in Basque for datives generally, and in Spanish for objects (DOM°, what should prevent that it be added to an object position in Basque? This is a free extension of a structural possibility. It triggers a narrowing in the set of interpretations that the DOM marked object gets in Basque. I expressed this by obligatory presence of [animate] in v.