

ACQUISITION OF FUNCTIONAL CATEGORIES IN CATALAN AND
FRENCH

**Defaults and Competition in the Acquisition of Functional Categories in
Catalan and French**

Lisa Davidson and Géraldine Legendre
Johns Hopkins University

To appear in:

*Selected Papers from the 2001 Linguistic Symposium on Romance
Languages (LSRL XXVI).*

Lisa Davidson and Géraldine Legendre
Department of Cognitive Science
Johns Hopkins University
3400 N. Charles St.
Baltimore, MD 21218
[davidson, legendre]@cogsci.jhu.edu

Defaults and Competition in the Acquisition of Functional Categories in Catalan and French

Lisa Davidson and Géraldine Legendre
Johns Hopkins University

1. Introduction*

One of the enduring topics in the study of language acquisition is children's seemingly optional use of non-finite forms as root verbs. This phenomenon, usually called the "optional infinitive" stage, is intriguing in part because it does not appear to occur in the acquisition of all languages. The typical claim is that children learning null-subject languages with rich inflection such as Italian, Catalan, and Spanish rarely use non-finite root forms (NRFs), or use them only at early stages and for a very short period of time (e.g. Guasti 1994, Torrens 1995). On the other hand, children learning obligatory subject languages, often with more limited inflectional paradigms, such as German, Swedish, Dutch, or English have higher proportions of NRFs that persist for a longer period of time (e.g. Weverink 1989, Wexler 1994, Phillips 1995).

Traditionally, accounts of the acquisition of tense and agreement have focused on the second set of languages—those which exhibit a high proportion of NRFs during acquisition. However, recent work on the acquisition of null subject/richly inflected languages suggests that "default"-like forms may not be limited to the NRFs found in the speech of children acquiring relatively poorly-inflected languages. Ferdinand (1996) observed that French speaking children overgeneralize the third person singular present indicative form. This has also been suggested for Spanish and Catalan by Grinstead (1998).

The acquisition of inflectional categories raises the question of whether tense and agreement should properly be considered as occurring simultaneously, or as two separate events that may differ in time course both within and across languages. According to Meisel (1994) the pattern and order of acquisition of tense and agreement is the same across all languages, but little rigorous cross-linguistic work has backed up that claim.

One further shortcoming of previous accounts of the optional infinitive stage is that they make no attempt to account for the actual proportions of finite and default forms observed in the child's speech. In this paper, we appeal to two contemporary principles of Optimality Theory (Prince and Smolensky 1993) to explain changing proportions in the speech of children acquiring French and Catalan. One concerns the initial state of a child's grammar and the proposal that faithfulness constraints (those which compel faithful production of the input) are ranked below markedness constraints (those which require outputs to conform to certain structural restrictions) (Demuth 1995, Gnanadesikan 1995, Smolensky 1996). During learning, a child compares her output to input from adult speakers and

changes her constraint ranking in an effort to minimize differences between her production and the adult one (Tesar and Smolensky 2000).

The other contemporary OT principle pertains to properties of the constraint rankings themselves. Based on the observation that children often exhibit alternations between correct and non-adult-like outputs, proposals using non-strictly ranked or “floating” constraints have been developed to explain how child speech and other types of variation contain stages where multiple grammars can coexist and different outputs can be optimal for a given input depending on which grammar is active (e.g. Anttila 1997, Boersma 1998, Boersma and Hayes 1999, Smolensky, Davidson and Jusczyk to appear). Floating constraints and partial rankings will play a central role in explaining the proportions of default and of tensed and/or agreeing forms in child French and Catalan.

Overall, we explore the similarities and differences in the patterns of acquisition of their native language by French and Catalan children. First, we examine the proportions of default forms found in the children utterances in both languages. We follow up on the original suggestion by Grinstead (1998) that third person singular present indicative forms may actually be default forms in Catalan and probe the data further. Next, we examine the time course of the acquisition of tense and agreement. Whereas French children begin to acquire tense first, Catalan child speech initially contains agreeing but not tensed utterances. Furthermore, while Catalan children exhibit a linear increase in the acquisition of both tense and agreement, French children exhibit a stage where there is competition for the realization of functional categories. Finally, a detailed analysis for the Catalan child Laura is developed as an example of how Optimality Theory can account for variation both within and across languages.

2. Empirical patterns in French and Catalan

2.1 Developmental stages

The data for both French (Champaud, Rondal, and Suppes, Smith, and Leveillé corpora) and Catalan (Serrà-Solé corpus) come from the CHILDES database (MacWhinney and Snow 1985). The French children include Grégoire (1;9-2;5), Stéphane (2;2-3;3) and Philippe (2;1-2;6). The Catalan children examined were Pep (1;6.23-3;3.18), Gisela (2;1.23-3;10.2), and Laura (1;9.7-3;5). The breakdown of the data into developmental stages was determined using the Predominant Length of Utterance (PLU) measure developed by Vainikka, Legendre and Todorova (1999). This measure was created as an alternative to age and MLU (Mean Length of Utterance) in an attempt to better characterize syntactic development. Briefly, the PLU measure identifies length of utterance and occurrence vs. absence of verbs as syntactic milestones (see appendix for details).

For the Catalan children, four stages of development are demonstrated, labeled here as 2b, 3b, 4b, and 4c. One child, Pep, also showed an intermediate stage that alternated between 3b and 4b which is

denoted as 3b/4b. The available data for the French children covers stages 3b, 3c, and 4c. For both the French and Catalan children, 4c represents an adult-like stage in the development of functional categories. Due to lack of space, the age of the children at each PLU stage and the number of utterances is not presented here. Detailed information can be found in Legendre et al. (1999) for French and Davidson (2001) for Catalan.

2.2 NRFs and defaults

For children in the “optional infinitive” stage of grammatical development, root infinitives are possible grammatical utterances that can coexist with finite forms (Wexler 1998). Here, we assume that not only infinitives, but rather any non-finite forms, may appear as root verbs in a child’s speech during this stage. Under this definition, NRFs include infinitives, bare participles, and bare gerunds. Examples of NRFs in child speech are given in (1) and (2).

- (1) French Cabinets ouvrir.
 Restroom open-INF
 “(I will) open the restroom (door)” (Grégoire 3b—1;9.28)
- (2) Catalan Tirat un nen a terra.
 throw-PARTIC a boy to ground
 “<I> threw a doll on the floor.” (Pep 3b—2;1.1)

For the French children, the number of NRFs at stage 3b is relatively high, and they decrease until stage 4c. Catalan children seem to use considerably fewer NRFs than the French children.

Table 1. Proportion of NRFs of all verbs for French

Child	Stage 3b	Stage 4b	Stage 4c
Grégoire	28% (83/297)	18% (51/287)	1% (7/711)
Stéphane	48% (51/106)	13% (27/205)	2% (3/152)
Philippe	—	22% (105/476)	6% (14/250)

Table 2. Proportion of NRFs of all verbs for Catalan

Child	Stage 2b	Stage 3b	Stage 3b/4b	Stage 4b	Stage 4c
Pep	20% (9/45)	10% (10/105)	3% (8/229)	0% (1/314)	0% (1/151)
Gisela	5% (1/22)	4% (2/45)	—	1% (2/334)	0% (1/294)
Laura	3% (1/34)	8% (17/217)	—	2% (7/293)	1% (2/176)

Of the Catalan children, only Pep’s utterances contain more than 10% NRFs at any stage, and these are mostly bare participles. At first glance, this data seem to suggest that children acquiring Catalan do not use NRFs. However, as suggested by Grinstead (1998), it may be that Catalan learners are still using some sort of default form. Grinstead’s original speculation was based on two observations: (1) children use a very high proportion of 3rd person present singulars (3S-PI), and (2) most of their

agreement errors involve substituting a 3S-PI form for a 1st person singular (1S), 2nd person singular (2S), or 3rd person plural (3P) target.

In an effort to quantify this intuition, we examined every 3rd person singular form to uncover those used as defaults. It is evident that in many cases, children both intend and correctly produce a 3S target. Consequently, a 3S form is only considered a 3S-PI default if it is clearly used when another form was intended. These can be either person agreement errors, number agreement errors, or errors resulting from the omission of the impersonal-*se* clitic. Examples of such cases are given in (3):

(3) a. Number agreement errors (Laura 3b—2;8.30)

Dad: I on són les boles? and where are-3P the balls? “And where are the balls?”	Lau: No hi és. not here is-3S “[It] is not here.”
---	---

b. Person agreement errors (Pep 3b—2;1.1)

Pep: Ai, no pot! EXCLAM, no can-3S “Ai, [he] can’t!”	Laia: Si que pots. yes that can-2S “Yes, you can”
--	---

c. Missing impersonal-*se* (Laura 3b—2;5.8)

Lau: Talla carn.
cut-3S meat
“Meat is cut” or “I’m cutting meat”

In principle, use of 3S forms in place of 1S could be a reference error rather than an agreement one. However, this does not appear to be the case for any of the children analyzed here. None of them ever overtly use their proper name with a 3S form to refer to themselves (in contrast to the child Guillem, from the same corpus, who frequently uses his own name with 3S verbs). In addition, correct use of first person agreement occurs simultaneously with this type of agreement errors in a given stage. Thus, we interpret use of a 3S form in place of 1S as an agreement error here, not a reference one.

Like NRFs, 3S-PI defaults are regarded as non-finite.¹ If NRFs and 3S-PI defaults are considered together, then it no longer seems as if Catalan learners’ verbal utterances are adult-like from the earliest stages. Comparing Table 3 with Table 1, we can see that the proportion of default forms in Catalan child speech is now more like that for French. A thorough analysis of child Catalan must include an account of this phenomenon.

Table 3. Proportion of default forms in Catalan child utterances

Child	Stage 2b	Stage 3b	Stage 3b/4b	Stage 4b	Stage 4c
Pep	44% (20/45)	12% (13/105)	6% (13/229)	0% (1/314)	0% (1/151)
Gisela	18% (4/22)	7% (3/45)	—	1% (3/334)	0% (1/294)
Laura	47% (16/34)	15% (33/217)	—	3% (10/293)	1% (2/176)

2.3 Development of tense and agreement

Following previous syntactic analyses of the representation of tense (TNS) and agreement (AGR), it is assumed that TNS and AGR each project a functional head (Pollock 1989). It has been proposed that in non-finite child utterances, either AgrP and TnsP can be missing from the representation (Schütze and Wexler 1996, Wexler 1998). We adapt and expand this set of possibilities to allow for the realization of TNS or AGR only, both TNS and AGR, or neither. In the case that either TNS or AGR is omitted from the utterance, default morphology for the missing feature will surface. For both French and Catalan, we consider default TNS to be the present indicative and default AGR to be 3rd person singular. This is exemplified in (4) with Catalan for a 1S, past tense input:

(4) Missing feature	Resulting surface form & features
TNS & AGR	<i>menjar/menjat/menja</i> (eat-INFIN./PARTIC./3S-PI)
AGR only	<i>ha menjat</i> (eat-3S, past)
TNS only	<i>menjo</i> (eat-1S, present)
Neither	<i>he menjat</i> (eat-1S, past)

Contra Schütze and Wexler (1996) we consider verbs containing either TNS or AGR or both to be finite; only those lacking both TNS and AGR are non-finite. According to Meisel (1994), acquisition of TNS and AGR occurs when a child productively uses more than one tense or agreement affix. Based on these proposals, the following criteria are used to determine the proportion of tense and agreement in child utterances:

- (5) a. A form is counted as *tensed* if it is not present (since present tense is ambiguous between default and correctly agreeing forms).
- b. A form is counted as *agreeing* if it is not 3S.
- c. *Non-default* 3rd singular present indicative forms are not counted as either tensed, agreeing or default since they have ambiguous TNS and AGR. These forms are excluded from further counts.

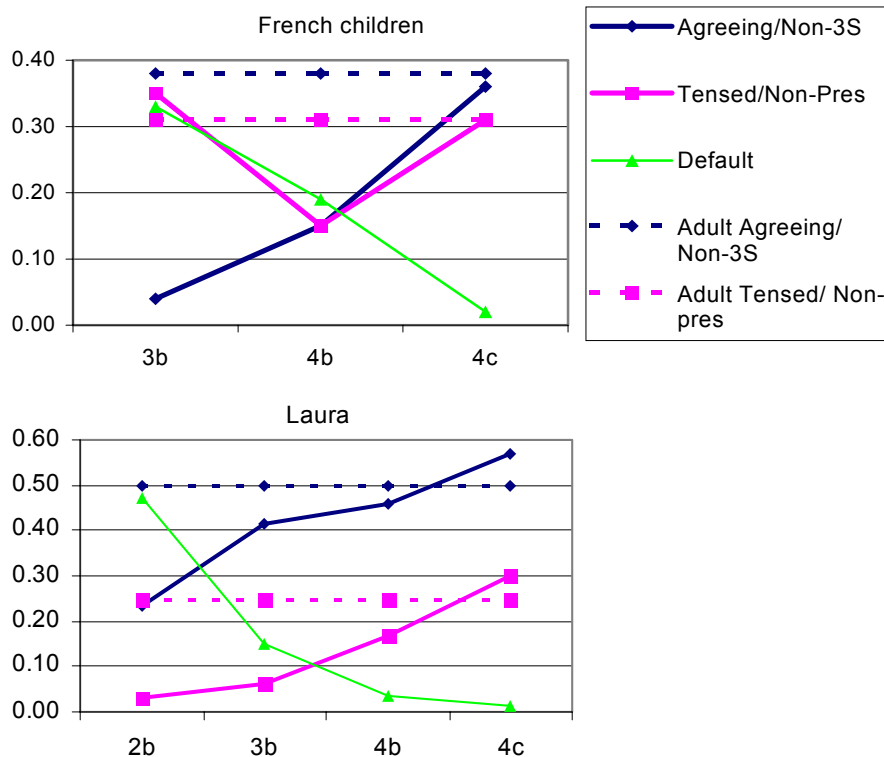
For French, it is assumed that subject clitics are overt realizations of agreement, which entails that French is a null-subject language (following e.g. Roberge 1990, Pierce 1992, Kaiser 1994, Ferdinand 1996, Legendre et al. 1999). With respect to the criteria in (5), a (regular) verb in French is counted as agreeing if it appears with a subject clitic.

In order to measure the child's syntactic progress, we can compare the proportions of tensed and agreeing forms in their speech according to the criteria in (5) with the proportions found for adults using the same counting method. Evidently, we are not claiming that adult 3S and present tense utterance are not agreeing or tensed; we are simply using the criteria in (5) as a metric for evaluating the progress of the children. Once the children have attained adult-like proportions of tense and agreement according to this metric, then we can assume that all of their verbal utterances are

syntactically adult-like. Both the child and adult proportions are shown graphically in Figure 1.

(Note: Because the French child patterns are so similar, their data have been collapsed. From here on, Laura's data will be presented to exemplify the Catalan pattern. While the proportions are slightly different, Laura's and Gisela's patterns of development are very similar. Pep's extra stage requires an additional level of analysis that is beyond the scope of this paper. See Davidson (2001) for a full analysis of each child.)

Figure 1. French vs. Catalan developmental patterns



The patterns in Figure 1 can be summarized as in (6):

(6) a. *French*

TNS: Tense starts out being used at adult-like levels (3b).

Tense use drops dramatically (4b).

Tense returns to adult-like levels (4c).

AGR: Agreement starts out essentially unused (3b).

Agreement increases (4b).

Agreement reaches adult-like levels (4c).

DEF: NRFs decrease

b. *Catalan*

TNS: Tense starts out essentially unused (2b)

Tense increases (3b, 4b)

Tense reaches adult-like levels (4c)

AGR: Agreement starts ~50% of adult (2b)

Agreement increases (3b, 4b)

Agreement reaches adult-like levels (4c)

DEF: Defaults decrease

From these patterns, it can be concluded that TNS and AGR show different courses of acquisition in both French and Catalan and that French and Catalan differ from one another. For Catalan, the proportion of AGR is initially greater than TNS, and the opposite is true for French. This may be due to the proportions of these features in the adult speech: for the French parent the proportion of overt TNS and AGR is roughly equal (31% TNS, 38% AGR), whereas there is twice as much AGR as TNS in Catalan parental verb forms (24% TNS, 50% AGR).

For the French children, we propose that the developmental pattern can be explained by assuming competition for a single functional projection (FP). By stage 3b, the child has only one FP, and it is used to realize TNS. At stage 4b, there is still only one FP, but the child now tries to realize both TNS and AGR. TNS drops, because TNS and AGR are now competing for one FP. Finally at Stage 4c, the child now has 2 FPs, and both TNS and AGR can be realized.

The patterns of development in Figure 1 clearly indicate that from stage to stage, the proportions of TNS and AGR in the children's utterances do not increase discretely, but rather continuously. Likewise, the use of defaults does not entirely disappear once children begin to use fully inflected forms. That acquisition of functional projections is not a discrete process has been recognized (Phillips 1995) but most detailed accounts do not even address the need to explain such variation (e.g. Pierce 1992, Rizzi 1993/94, Vainikka 1993/94, Meisel 1994, Wexler 1994, 1998). We make the case next that OT provides the necessary constructs to account for the proportions of TNS, AGR and Default forms at different acquisition stages.

3. Accounting for variability in Optimality Theory

3.1 Floating constraints and partial rankings in child grammars

As discussed in Section 2, child verbal utterances are not always specified for TNS and/or AGR. According to Schütze and Wexler (1996) and Wexler (1998), this may be because realizing both TNS and AGR might be too complicated for children, who compensate by deleting one of either TNS or AGR features. In other words, the child is facing conflicting constraints. Since resolving constraint conflict is a fundamental aspect of Optimality Theory, the restrictions that the child faces can be encoded in the following constraints:

- | | |
|--|------------------------------|
| (7) Economy of structure constraints | (8) Faithfulness constraints |
| *F: No functional heads | PARSET: Parse Tense |
| *F ² : No pairs of functional heads | PARSEA: Parse Agreement |

Note that *F² is the local conjunction of *F with itself, and that these constraints are part of a power hierarchy: *F² >> *F in every grammar (Smolensky 1997). PARSET and PARSEA are not universally fixed.

Four candidate structures are relevant to the OT analysis. It is assumed that the *input* to each evaluation contains TNS and AGR features, and that the realization or absence of these features is dependent on the ranking of the constraints in the grammar. Thus, if both markedness constraints are ranked above both faithfulness constraints, neither TNS nor AGR will be realized. The relevant candidate structures for an input with [1S] and [past] features and the constraints they violate and satisfy are shown in (9):

(9) Possible structures (candidates) evaluated for “optimality”:

- | | | |
|----|--|--|
| a. | VP | example: <i>cantar</i> (NRF)
violates: PARSEA, PARSET
satisfies: *F, *F ² |
| b. | <pre> TP / \ T[PAST] VP </pre> | example: <i>ha cantat</i> (3S, past)
violates: PARSEA, *F
satisfies: PARSET, *F ² |
| c. | <pre> AgrP / \ Agr[1S] VP </pre> | example: <i>canto</i> (1S, present)
violates: PARSET, *F
satisfies: PARSEA, *F ² |
| d. | <pre> AgrP / \ Agr[1S] TP / \ T[Past] VP </pre> | example: <i>he cantat</i> (1S, past)
violates: *F, *F ²
satisfies: PARSEA, PARSET |

Usually in Optimality Theory, it is assumed that there is strict domination of constraints: for every two constraints C₁ and C₂, either C₁ outranks C₂ or C₂ outranks C₁. If this were true in child grammars, then we would not see continuously changing proportions of TNS and AGR in child speech. However, if we consider that two or more grammars may exist simultaneously, we can hypothesize that a child may “choose” between them, thus producing tensed and/or agreeing forms sometimes but not always. This type of variation can be captured through partial rankings, which give rise to different proportions of utterances inflected for TNS and/or AGR. A hypothetical partial ranking is shown schematically in (10):

- (10) a. Set of rankings:
- i. *F² >> *F >> PARSET winning candidate: untensed verb
 - ii. *F² >> PARSET >> *F winning candidate: tensed verb
- b. Partial ranking that will make both grammars possible



In the example in (10), the floating constraint PARSET covers a range of the ranking space. Prior to each evaluation, PARSET is fixed in a random place in that range, so that either (strict) ranking (10)a.i. or (10)a.ii. results. One logical possibility is that any location in PARSET's range has an equal probability of being the place where PARSET is fixed in a given evaluation. In the preceding example, this means that an *untensed verb* is realized 50% of the time while the *tensed verb* is realized 50% of the time. This is in fact what seems to happen with the French children. For a detailed OT analysis of the French data, see Legendre et al. (1999).

The other logical possibility is that the floating range of a constraint is not equiprobable, such that children will have a bias in terms of producing more or fewer forms with TNS and/or AGR. This is the case for Catalan. Space considerations prevent us from discussing each child individually, but Laura is representative of all three Catalan children. An in-depth OT analysis of Pep and Gisela can be found in Davidson (2001).

3.2 Theoretical advantages of an OT approach

Previous accounts of syntactic acquisition within the Principles and Parameters framework have generally promoted one of two approaches: Strong Continuity/Full Competence or Weak Continuity/Minimal Trees. According to the former viewpoint, adult-like phrase structure is available to the child from the beginning of syntactic acquisition. Any discrepancies between child and adult syntax are attributed to external factors such as memory limitations or child-specific restrictions on particular grammatical structures (Boser, Lust, Santelmann and Whitman 1992, Wexler 1994, 1998). Weak Continuity, on the other hand, assumes that child and adult grammars are not the same. A child's syntactic representation begins with only the most basic projection—a VP—and then develops gradually (Guilfoyle and Noonan 1992, Vainikka 1993/94, Radford 1996).

The advantage of the optimality-theoretic partial ranking analysis is that it combines elements of both the Strong and Weak Continuity approaches, which were previously viewed as incompatible. The properties of these two approaches incorporated by the partial ranking analysis are summarized as follows (from Legendre et al. 1999):

Strong Continuity Properties shared by the partial ordering analysis

- The input to the syntactic component is adult-like (functional categories are present in the grammar from the beginning of syntactic acquisition)
- The syntactic constraints are adult-like (no constraints are specific to child language)
- Variation between stages of a child's grammar is adult-like (constraint reranking accounts for different child grammars just as it does for historical or sociolinguistic variation in adult grammars.)

Weak Continuity Properties shared by the partial ordering analysis

- Each stage of development corresponds to a different grammar (the differing grammars arise from constraint reranking during development)
- Minimal amount of phrase structure is posited (this economy of structure arises from interaction between markedness and faithfulness constraints)
- There exist non-adult-like grammars at early stages of development (such grammars arise from a general tendency of faithfulness constraints to have an initial low ranking).

3.2 Case study from Catalan: Laura

As determined by the criteria in (5) above, the child's syntactic progress can be calculated as a function of the proportions of tensed and agreeing forms in the adult speech. For example, if 24% of an adult's verbal utterances contain non-present tense, and 12% of all child verbal utterances do, then we say that the child is reaching her target for tense only 50% of the time. The proportions of TNS, AGR and both TNS&AGR in a child's speech are calculated using the following formulas:

$$(11) \quad \text{Prop}(\text{child AGR}) = \frac{\text{child non-3S}}{\text{adult non-3S}} \quad \text{Prop}(\text{child TNS}) = \frac{\text{child non-present}}{\text{adult non-present}}$$

$$\text{Prop}(\text{AGR\&TNS in child speech}) = \frac{\text{child non-3S, non-present}}{\text{adult non-3S, non-present}}$$

Using this measure for determining how frequently a child uses TNS and/or AGR with respect to her ultimate goal of attaining the adult proportions, Laura's data can be examined stage-by-stage. Her proportions for stage 2b are given in Table 4. (Note: In the following tables, "Proportion" refers to the proportion of TNS and AGR in her utterances compared to the adult target.)

Table 4. Laura, Stage 2b	AGR (NON-3S)	TNS (NON PRESENT)	AGR&TNS (NON3S, NONPRES)
Laura	24% (8/34)	3% (1/34)	0% (0/34)
Laura's Adults	50% (341/682)	25% (168/682)	13% (86/682)
Proportion	48%	12%	0%

From this data, we see that 48% of Laura’s utterances are agreeing compared to the adult target, and 12% are tensed. She has no utterances that are both tensed and agreeing at this stage (e.g. first person, past tense). In OT terms, this entails that PARSEA floats above (at least) *F 48% of the time, and PARSET does so 12% of the time. Note that by the definition of probability, this means that PARSEA is ranked below *F (100%–48%) = 52% of the time, and PARSET is ranked below *F (100%–12%) = 88% of the time. Her partial ranking for this stage is given in (12):

- (12) Laura, 2b—Rankings (conservatively assuming that neither faithfulness constraint is ever above *F², since none of her utterances are both tensed *and* agreeing at this stage)

*F ²	>>	*F	
	48%	52%	PARSEA
	12%	88%	PARSET

The schematic in (12) translates into six competing grammars at stage 2b, shown in (13). However, the likelihood of each grammar being used at any given time is not equiprobable, but biased toward those which realize AGR.

- (13) *F² >> *F >> PARSEA >> PARSET *F² >> PARSET >> *F >> PARSEA
 *F² >> *F >> PARSET >> PARSEA *F² >> PARSEA >> PARSET >> *F
 *F² >> PARSEA >> *F >> PARSET *F² >> PARSET >> PARSEA >> *F

One consequence of this analysis is that we can confirm the metric developed to determine the default. By simply knowing Proportion(TNS) and Proportion(AGR), the default can be predicted and checked against the observed. Recall that whenever PARSEA and PARSET are floating together at the bottom of the hierarchy, a default form results. The rankings in (14) will yield default forms:

- (14) *F² >> *F >> PARSEA >> PARSET *or* *F² >> *F >> PARSET >> PARSEA

The proportion of default forms can be predicted by determining how often PARSEA and PARSET will co-occur at the bottom of the hierarchy. We already know that for Laura at Stage 2b, PARSEA is below *F 52% of the time, and PARSET is below *F 88% of the time. Multiplying these numbers together gives us the *joint probability* that both PARSEA and PARSET will be floating together below *F. This number tells us what the *predicted* default is, as shown in (15):

- (15) Predicted default = P(1-Proportion(AGR)) × P(1-Proportion(TNS))
 = 52% × 88% = 47%

The predicted default, 47%, is the same as the observed proportion of default forms in Laura’s speech, as reported in Table 3.

One assumption of this theory is that the children will never “go back” on their rankings; in other words, the proportions of TNS and AGR in their speech should be continually increasing. For Laura at stage 3b, this is indeed true, as illustrated in Table 5:

Table 5. Laura, Stage 3b	AGR (NON-3S)	TNS (NON PRESENT)	AGR&TNS (NON3S, NONPRES)
Laura	41% (90/217)	6% (13/217)	4% (9/217)
Laura’s Adults	50% (341/682)	25% (168/682)	13% (86/682)
Proportions	83%	24%	33%

At this stage, Laura has forms that are both tensed and agreeing. This entails that both PARSET and PARSEA must be floating above *F^{2,2}.

(16) Laura, 3b—Rankings

*F ²	>>	*F	
83%		17%	PARSEA
24%		76%	PARSET

The predicted default for stage 3b can again be compared to the observed proportion, again from Table 3:

(17) Default = 17%×76% = 13% (observed=15%, $\chi^2(1)=.88, p > .25$)

The data for Laura at stage 4b is much like that for 3b, except that the proportions of AGR, TNS, and TNS&AGR are all closer to 100%. This is shown in Table 6 and (18):

Table 6. Laura, Stage 4b	AGR (NON-3S)	TNS (NON PRESENT)	AGR&TNS (NON3S, NONPRES)
Laura	46% (134/293)	17% (49/293)	8% (24/293)
Adult	50% (341/682)	25% (168/682)	13% (86/682)
Proportions	91%	71%	65%

(18) Laura 4b—Rankings

*F ²	>>	*F	
91%		9%	PARSEA
71%		29%	PARSET

Again, the predicted default matches the observed default:

(19) Default = 9%×29% = 2.6% (observed=3% $\chi^2(1)=.73, p > .25$)

By stage 4c, Laura’s proportions of TNS and AGR are essentially adult-like. We assume that both PARSEA and PARSET are ranked above *F² by this stage, as shown in Table 7 and (20).

Table 7. Laura, stage 4c	AGR (NON-3S)	TNS (NON PRESENT)	AGR&TNS (NON3S, NONPRES)
Laura	57% (100/176)	30% (53/176)	19% (33/176)
Adult	50% (341/682)	25% (168/682)	13% (86/682)

(20) Laura 4b—Rankings
 *F² >> *F

100%	PARSEA
100%	PARSET

The final ranking for Laura—the one which she shares with adults—is PARSEA, PARSET >> *F² >> *F.

4. Conclusion

In this paper, we have presented three main findings. First, while French has always been considered an optional infinitive language, it has been assumed that Catalan is not. When the phenomenon is couched in more general terms, namely the use of default forms in acquisition, Catalan looks very much like French. Catalan learners too use a default form which lacks tense and agreement: the third person singular present indicative.

Second, we have demonstrated that tense and agreement follow distinct patterns of development, both within and across languages. Whereas French learners begin acquiring TNS first, Catalan learners acquire AGR first. Furthermore, the development of TNS and AGR is essentially linear in Catalan while French learners show evidence of syntactic competition in an intermediate stage at which one functional projection is available, and TNS and AGR compete for realization.

Third, an optimality-theoretic analysis with partial constraint ranking can explain how varying proportions of tensed, agreeing, and default utterances arise. The analysis demonstrates how a child progresses from stage to stage until s/he has fully acquired tense and agreement.

Notes

* We would like to thank Paul Hagstrom, Marina Todorova, Anne Vainikka, Matt Goldrick, and the LSRL XXXI audience for their significant contributions to this work. All errors are our own. This research was carried out under NSF (LIS) grant # 9720412.

¹ Note that the realization of any given non-finite form may be dependent on different types of aspectual content that the child is intending to express (Bel i Gaya 1998, Hyams 2000). This is not a problem for Optimality Theory, since the relevant aspectual features would be present in the input, but absent on the surface due to relatively high ranked markedness constraints dispreferring their realization.

² Note that while the presence of forms that are both agreeing and tensed entails that the faithfulness constraints must float above *F² at least part of the time, the proportion of the time that they can be found only above *F versus above *F² is underdetermined. This is because there are often several different rerankings and accompanying frequencies which could give rise to the observed Proportion(AGR&TNS). For a more detailed explanation, see Davidson (2001) and Davidson and Goldrick (2001).

References

- Anttila, A. 1997. Variation in Finnish Phonology and Morphology. Ph.D. dissertation, Stanford University.
- Bel i Gaya, A. 1998. Teoria lingüística i adquisició de llenguatge: Anàlisi comparada dels trets morfològics en català i castellà. Ph.D. dissertation, Universitat Autònoma de Barcelona.
- Boersma, P. 1998. Functional Phonology: Formalizing the interactions between articulatory and perceptual drives. The Hague: Holland Academic Graphics.
- Boersma, P. and B. Hayes. 1999. Empirical tests of the Gradual Learning Algorithm. Rutgers Optimality Archive ROA-348, University of Amsterdam and UCLA. <http://rucss.rutgers.edu/roa.html>
- Boser, K., B. Lust, L. Santelmann and J. Whitman. 1992. "The syntax of CP and V2 in early child German: The Strong Continuity Hypothesis". In K. Broderick, ed., Proceedings of NELS 23, 51-65. Amherst, Mass: UMass GLSA
- Davidson, L. 2001. Defaults, frequencies and variation in the acquisition of Catalan. ms., Johns Hopkins University.
- Davidson, L. and M. Goldrick. 2001. Uncovering default verbs in the acquisition of Catalan. Talk presented at the 4th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages, University of Illinois, October 11-14, 2001.
- Demuth, K. 1995. "Markedness and the development of prosodic structure". In J. Beckman, ed., Proceedings of the Northeastern Linguistics Society, 13-25. Amherst, Mass.: UMass GLSA
- Ferdinand, A. 1996. The Development of Functional Categories: The Acquisition of the Subject in French. The Hague: Holland Academic Graphics.
- Gnanadesikan, A. 1995. Markedness and faithfulness constraints in child phonology. Rutgers Optimality Archive ROA-67, University of Massachusetts, Amherst. <http://rucss.rutgers.edu/roa.html>
- Grinstead, J. 1998. Subjects, Sentential Negation and Imperatives in Child Spanish and Catalan. Ph.D. dissertation, UCLA.
- Guasti, M. 1994. "Verb syntax in Italian child grammar: Finite and non-finite verbs". Language Acquisition 3 (1). 1-40.
- Guilfoyle, E. and M. Noonan. 1992. "Functional categories and language acquisition". Canadian Journal of Linguistics 37. 241-272.

- Hyams, N. 2000. Now you hear it, now you don't: The nature of optionality in child grammars. Plenary talk presented at the Boston University Conference on Language Development, November 4, 2000.
- Kaiser, G. 1994. "More about INFL-lection and agreement: The acquisition of clitic pronouns in French". In J. Meisel, ed., *Bilingual First Language Acquisition: French and German Development*. Amsterdam: John Benjamins
- Legendre, G., P. Hagstrom, A. Vainikka and M. Todorova. 1999. Partial constraint ordering in child French syntax. Technical Report, Johns Hopkins University.
- MacWhinney, B. and C. Snow. 1985. "The Child Language Data Exchange System". *Journal of Child Language* 12. 271-296.
- Meisel, J. 1994. "Getting FAT: Finiteness, agreement, and tense in early grammars". In J. Meisel, ed., *Bilingual First Language Acquisition: French and German Grammatical Development*. Philadelphia: John Benjamins
- Phillips, C. 1995. "Syntax at age two: Cross-linguistic differences". *MIT Working Papers in Linguistics* 26. 37-93.
- Pierce, A. 1992. *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars*. Dordrecht: Kluwer.
- Pollock, J.-Y. 1989. "Verb movement, Universal Grammar, and the structure of IP". *Linguistic Inquiry* 20. 365-424.
- Prince, A. and P. Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical Report Rutgers University and Technical Report, University of Colorado, To appear, MIT Press.
- Radford, A. 1996. "Towards a structure building model of acquisition". In H. Clahsen, ed., *Generative Perspectives on Language Acquisition*, 43-90. Philadelphia: John Benjamins
- Rizzi, L. 1993/94. "Some notes on linguistic theory and language development: The case of root infinitives". *Linguistic Inquiry* 17. 371-393.
- Roberge, Y. 1990. *The Syntactic Recoverability of Null Arguments*. McGill: Queen's University Press.
- Schütze, C. and K. Wexler. 1996. "Subject case licensing and English root infinitives". In A. Stringfellow, D. Cahan-Amitay, E. Hughes and A. Zukowski, eds., *Proceedings of the 20th Annual Boston University Conference on Language Development*, 670-681. Somerville, Mass.: Cascadilla Press
- Smolensky, P. 1996. The initial state and 'richness of the base' in Optimality Theory. Technical Report JHU-CogSci-96-4 and Rutgers Optimality Archive ROA-154, Cognitive Science Department, Johns Hopkins University, Baltimore, MD. <http://ruccs.rutgers.edu/roa.html>

- Smolensky, P. 1997. Constraint interaction in generative grammar II: Local Conjunction (or, Random rules in Universal Grammar). Paper presented at the Hopkins Optimality Theory Conference/University of Maryland Mayfest, Baltimore, MD.
- Smolensky, P., L. Davidson and P. Jusczyk. to appear. "Optimality in language acquisition I: Phonology". In P. Smolensky and G. Legendre, eds., *The Harmonic Mind: From Neural Computation to Optimality Theoretic Grammar*. Cambridge, MA: Blackwell
- Tesar, B. and P. Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge, Mass.: MIT Press.
- Torrens, V. 1995. *L'adquisició de la sintaxi en català i castellà: La categoria funcional de flexió*. Ph.D. dissertation, Universitat de Barcelona.
- Vainikka, A. 1993/94. "Case in the development of English syntax". *Language Acquisition* 3. 257-324.
- Vainikka, A., G. Legendre and M. Todorova. 1999. PLU-Stages: an independent measure of early syntactic development. Technical Report JHU-CogSci-99-10, Johns Hopkins University.
- Weverink, M. 1989. *The subject in relation to inflection in child language*. M.A. thesis, University of Utrecht.
- Wexler, K. 1994. "Optional infinitives, head movement, and economy of derivation". In N. Hornstein and D. Lightfoot, eds., *Verb Movement*, 305-350. Cambridge: Cambridge University Press
- Wexler, K. 1998. "Very early parameter setting and the Unique Checking Constraint". *Lingua* 106. 23-79.

Appendix. Developmental stages

Definitions of PLU Stages (from Vainikka et al. 1999)

Stage 1: Predominantly one-word stage

- Almost all utterances (90%) are of the one-word sentence type

Stage 2: Intermediate stage between one-word and two-word stage

- The one-word sentence type is still very common (60%-89% of the utterances are of the one word-type)

Stage 3: "Two-word" stage

- The one-word sentence type no longer clearly predominates (i.e. fewer than 60% of all utterances are one-word utterances)
- The multiword sentence type is not the most common one

Stage 4: Predominantly multiword stage

- The multiword sentence type is the most common one

Secondary PLU stages (from Vainikka et al. 1999)

Secondary stage a: at most 10% of all utterances contain a verb

Secondary stage b: 11%-60% of all utterances contain a verb

Secondary stage c: more than 60% of all utterances contain a verb