

## Acquisition and Creolization of Condition C "violations" in Kadiwéu and Portuguese

Filomena Sandalo  
MIT

Peter Gordon  
University of Pittsburgh

### 1. Introduction

Within the principles and parameters approach to language acquisition, the focus of much research is on how and when children acquire the relevant principles that constrain the possible structures and interpretations within the language they are acquiring. The acquisition of constraints is generally considered to be underdetermined by the language input that children experience. Since constraints are negative in nature and the child receives little or no negative evidence that would be sufficient to specify the form of such constraints, they are considered to be a prime candidates for the innate component of the grammar.

Empirical research in this area has tended to take one of two paths. In some cases, researchers have demonstrated that constraints are present in children who are at the very beginning of acquiring the relevant structures, hence indicating the innateness of those constraints (Crain, 1991; Gordon, 1985). Alternatively, it might be found that such constraints are not present early on, and it is therefore necessary to develop accounts based on learning, maturation or the resetting of parameters.

In the present paper, we want to consider the issue of how constraints, whether acquired or innate, interact with the structures that undergird them. In particular, we will report on a study of children who are acquiring two languages, Kadiwéu and Portuguese, which differ fundamentally in the structures that specify the conditions for constraints on binding. In particular, while one of the languages, Portuguese, follows the universal constraint based on Condition C, Kadiwéu, on the other hand, shows apparent violations of this constraint. In this case, then, children are acquiring two languages that differ fundamentally in the interpretation of coreference constraints. In the present paper, we will present evidence that the Portuguese acquired by Kadiwéu children is creolized with respect to condition C. That is, they show judgments in Portuguese that reflect Kadiwéu binding conditions. If this is an accurate characterization, then it represents a significant insight that relatively covert properties can be infected from one language to another.

Condition C prohibits coreference of an R-expression with a c-commanding nominal element, thus prohibiting coreference of *he* and *John* in sentences like:

( 1 ) \*He<sub>i</sub> said that John<sub>i</sub> washed the dishes

Notice that when a language allows pro-drop, as in Portuguese, the empty category

<e> still obeys condition C in prohibiting the coreference of the empty category with the R-expression (*John*):

( 2 ) \*<e<sub>i</sub>> falou que Joao<sub>i</sub> lavou os pratos

Kadiwéu is typical of certain "polysynthetic" languages which are characterized by extensive dropping of nominal phrases leaving structures in which empty categories can be coreferent with nominal phrases in lower clauses in apparent violation of condition C, as in (3) and (4).

( 3 ) me:                      Joao    me            iwilegi                      Ginodi.  
*y-me:n*                      *Joao*    *me*            *y-wilegi*                      *Gino-di*  
 3sg.SUBJ-say    John    COMP    3sg.SUBJ-wash            plate-pl  
 '<e><sub>i</sub> said that the John<sub>i</sub> washed the dishes.'

( 4 ) yowo:Godi                      me            yema:                      Joao.  
*y-owo:-God*                      *me*            *y-ema:n*                      *Joao*  
 3sg.SUBJ-think-think    COMP    3sg.SUBJ-want            John  
 '<e><sub>i</sub> knows that e<sub>j</sub> loves John<sub>i</sub>.'

The questions that we will address in this paper are:

- Is condition C actually violated in Kadiwéu?
- How do children acquire the appropriate structures to determine binding conditions?
- Are there default parameters for binding?
- How do Kadiwéu children acquire Portuguese binding?
- How do Portuguese and Kadiwéu affect each other in this environment?
- How does Kadiwéu children's Portuguese differ from Brazilian children's Portuguese for binding?
- Are there asymmetries between Subject and Object?

## 2. Binding in Polysynthetic Languages

The first attempt to account for polysynthetic languages within the framework of generative theory was Hale's (1983) classification of these languages as "nonconfigurational", having the properties: free constituent order, pervasive dropping of NPs, and the existence of discontinuous constituents. Let us consider three proposals that attempt to explain why such languages might allow condition C violations.

Jelinek 1984 explains the properties of nonconfigurational languages by proposing that pronominal clitics and affixes are arguments; NPs are adjuncts, and therefore they can assume free order or be omitted. In the case of (3), for *y-wilegi* (*wash*), the clitic *y-* would be the subject, whereas the NP *Joao* would be an adjunct,

and not part of the argument structure of the verb. Since *Joao* is not an argument and is not c-commanded by the Subject inflection of the matrix verb, it can be coreferent with that subject without violating Condition C. This proposal has not been universally accepted primarily because the existence of inflectional morphemes functioning as arguments challenges the idea of a fixed phrase structure over which syntactic relationships such as subject and object can be defined.

An alternative analysis is presented by Baker (1995), who argues that polysynthetic languages are obligatory pro-drop languages where empty *pro* occupies the structural argument positions of the verb. Again, the Condition C "violations" that characterize polysynthetic languages follow from the fact that NPs are adjuncts. If an NP is an adjunct, the empty *pro* in argument position does not c-command it and Condition C again is not violated. This proposal, however, makes wrong predictions for Kadiwéu. If the verbal arguments were *pro* in Kadiwéu, one would expect that coreference between a main-clause argument and a nominal phrase in the complement clause would be impossible, since *pro* c-commands any of the NPs inside that complement clause. This is not the case. A pronominal subject in the main clause can be coreferent with a NP inside a complement clause in Kadiwéu, as in (3) and (4).

One could question whether the embedded clauses in (3) and (4) are indeed complement clauses. If they were adjunct clauses, Baker's analysis could be maintained for Kadiwéu because the coindexation of an argument in a main clause with a noun phrase in an adjunct clause does not violate Condition C (e.g. *I saw her before Mary died*). However, the analysis of the embedded sentences above as complement clauses is supported by the fact extraction from inside these sentence is allowed, as shown in 5. Example 6 shows that extraction from adjunct clauses is impossible.

( 5 )      igeimei                  me                  yema:                  me  
                  wh-place                  COMP                  3sg.SUBJ-want                  COMP

                 inoqe    Joao                  lod:a:jo?  
                  3sg.SUBJ-break-valency                  John                  3POSS-knife  
                  'Where<sub>i</sub> does he want to break John's knife t<sub>i</sub>?'

( 6 )      \*igeimei                  dapa:we    le:Godi                  Maria  
                  wh-place                  3sg.SUBJ-theme-scream                  because                  Mary

                 yaqadi                          napalwaGa?  
                  3SUBJ-find                  clay-pl  
                  \*'Where<sub>i</sub> did he scream because Mary found clay t<sub>i</sub>?'

In a third proposal, Chomsky (1995) suggests that NPs in polysynthetic languages are arguments of the verb, but that they are obligatorily moved into the non-argument (A-bar) specifier position of CP, thus leaving traces in argument

positions. Since it is the trace rather than the overt NP that is the argument, the latter is not c-commanded, and hence does not violate condition C. Kadiwéu appears to support this third position (see Sandalo, 1997, for discussion).

In the present study we were interested in whether Kadiwéu children show evidence of condition C violations in Kadiwéu and Portuguese. There are approximately 1,500 speakers of Kadiwéu living in the state of Matto Grosso do Sul in the southwest of Brazil. They occupy a large reservation of 500,000 hectares primarily raising cattle. There is a central village containing a school with small dispersed communities throughout the reservation. A history of violent conflicts with outsiders has led to some isolation. It is only in the last 15 years that there has been extensive contact with Brazilian culture. In this time, there have been several outside marriages, and some Kadiwéus have received higher education outside of the reservation. In the case of mixed marriages to non-Kadiwéu women, the language of the household is Portuguese. Children speak Kadiwéu to each other, and Portuguese is a second language to all adult native Kadiwéus.

### 3. The Experiment

Within the Kadiwéu tribe, we tested 31 Kadiwéu-speaking, and 18 Portuguese-speaking children. They ranged in age from 4 to 14 years. In addition, we tested 30 Brazilian children near Sao Paulo, aged 9 to 10. This age group was equivalent to the older Kadiwéu children with whom comparisons will be made in the analyses. Children were shown a picture and told a short story providing the context for sentences (7) (Subject coreference) and (8) (Object coreference). In each picture, there was a boy who was mentioned in the sentence, and another person, who could be the referent for the empty subject of the main verb "said". After hearing either (7) or (8), children were asked who said it?

#### Story 1

*Elyodo Joao owidi libaqetedi. Oda domaGa yema: Joao iwilegi lanodi. Oda me:ta Joao iwilegi GoniGinodi. NiGidiaGidi oda nGajo elyodo Joao jGigo iwi dantaGa igodi me iwilegi Ginodi.*

'John's mother had a lot of work. She wanted to ask John to do the dishes. Then she said: John, wash our dishes. Later John's mother came to see whether he had already washed the dishes.'

(7)      me:                      Joao      me                      iwilegi.  
             3sg.SUBJ-say    John      COMP              3sg.SUBJ-wash  
             '<e> said that the John washed the dishes.'

#### Story 2:

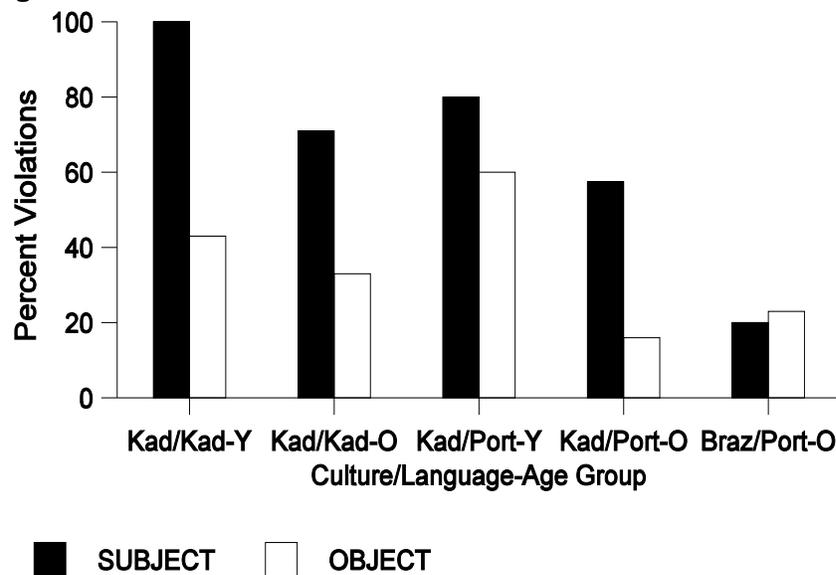
*Ika noqo ika Paulo yawaligeGe aka dom:o:jya, niGin:a naGa n:ad:i Jose me icomaGatike ika lad:igod:i. NGaka laqata oda Paulo ja niqoGeti nidom:ojya. Jose enitini.*

'One day Paul was driving a car, when he saw Joseph crossing the street. At this moment, Paul stopped the car. Joseph fell.'

( 8 )      me:                      me                      yamaGati                      dom:ojia.  
 3sg.SUBJ-say      COMP      3sg.SUBJ-hit-valency      car  
 '<e> said that the car hit Joseph.'

The data are summarized in Figure 1 separating the data for Younger (Age 4 to 7) and Older (Age 8 to 14) children. The graphs indicate the percentage of children who showed condition C violations (coindexing the subject of *said* with the proper name) for the Subject and Object conditions. The first striking fact about these data is that there is a clear subject-object Asymmetry. 77% of the Kadiwéu-speaking children allowed coreference with the embedded NP when it was the subject, and only 35% when it was the object. Surprisingly, this pattern also held for the Kadiwéu children who spoke Portuguese (67% subject, 30% object). Note that younger children allowed 100% of Condition C violations with subjects. The Brazilian children showed no such asymmetry between subject and object for Portuguese (20% subject, 24% object).

**Figure 1. SUBJ vs. OBJ Condition C Violations**



These results suggest that children do readily allow apparent violations of condition C in Kadiwéu. However, it also appears that the pattern of binding generalizes to the dialect of Portuguese spoken by the children in the tribe, suggesting a creolization of Portuguese with respect to binding constraints.



interact with Topic (old information) and Focus (new information):

$S_T VO_F \quad VO_T S_F \quad S_T O_F V \quad O_T S_F V \quad O_T V S_F \quad V S_F O_T$

This interaction of word order with Topic/Focus in the main clause indicates that there is NP dislocation into an A-bar position. According to Rizzi's (1996) analysis, Topics are located inside of CP. Support for the claim that topics are inside of CP in Kadiwéu comes from embedded clauses. Topics must precede the complementizer *naGa* in subordinate clauses.

Who is building this house? (i.e. talking about the house)

- (11) Maria yatemati di:m:igi naGa Pedro yoe.  
Mary 3SUBJ-say house COMP Peter 3SUBJ-make  
'Mary said that Peter is building the house.'

What is Peter building? (i.e. talking about Peter)

- (12) Maria yatemati Pedro naGa di:m:igi yoe.  
Mary 3SUBJ-say Peter COMP house 3SUBJ-make  
'Mary said that Peter is building the house.'

(see Sandalo 1997 for further discussion and evidence that focus NPs are also inside of CP).

**2. Verb Morphology.** Verbs are very highly inflected in polysynthetic languages. In Kadiwéu, there are 15 positions in the verb structure including morphemes including those for quantifiers, anaphors, and internal argument licensors (semantic role markers):

**3. Absence of non-referential quantified NPs.** Polysynthetic languages are characterized by lacking anaphoric nouns and quantifiers. Baker (1995) shows that the if NPs are always outside of IP, this correctly predicts that anaphoric expressions such as *himself/herself/themselves* and referential quantifiers such as *each other* should be nonexistent in polysynthetic languages. This is because their presence in such languages would lead to a violation of Condition A of the Binding Theory, which states that an anaphoric element must be bound by a c-commanding antecedent in argument position. If there are no NPs in argument position, then Condition cannot be met. Reflexives and reciprocals are thus expressed by verbal morphology. Baker also notes that the proposal that NPs are never inside of the IP predicts that polysynthetic languages should lack nonreferential quantified NPs (e.g. *everyone, everything, nobody, nothing*). Rizzi (1986) observes that quantified NPs cannot be dislocated in Italian (e.g. *\*Nessuno, lo conosco in questa città*) because pronominal clitics in the domain of a quantifier can only be interpreted as variables bound by that quantifier. But when a pronominal element is treated as a variable, the sentence is ruled out as an instance of vacuous quantification. If every NP is outside

of IP, it follows that quantifiers must be impossible in polysynthetic languages because the verb is obligatorily marked by pronominal clitics. In Kadiwéu quantifier notions are also expressed via verbal morphology.

**4. Verbs are derived.** In order to license grammatical internal arguments (i.e. transitive, ditransitive, and unaccusative clauses), Kadiwéu predicates require the presence of semantic role morphemes. Kadiwéu has six semantic role markers: *-d:-* 'theme', *-gi* 'goal', *-wa ~ -ma* 'dative', *-dom ~ -ma ~ -lo* 'benefactive', *-k* 'allative', and *lokom* 'adessive'. Examples (15) and (16) show that a stem functions as a noun or a transitive verb depending on whether the semantic role markers are present or not. In (15) they are not present and the stems function as a noun; that is, syntactic arguments cannot be added in spite of the fact that the valency suffix *-Gen:* '[+ internal argument]' is present and in spite of the fact that a Kadiwéu speaker understands that somebody is pierced. In (16), however, the semantic role *-d:-* 'theme' is present and the stems function as a transitive clause; that is, there is a grammatical internal argument, '1pl.OBJ'.

(15)    *lapwaGen:ig:i.*  
          *l:-apwa-Gen:-nig:i*  
          3POSS-hole-[+ internal argument]-m.dim  
          'His bodyguard.' (the one who becomes pierced)

(16)    *God:apwaGe*  
          *Go-d:-apwa-Gen:*  
          1pl.OBJ-theme-hole-[+ internal argument]  
          'We are challenged.'

Grimshaw (1990) proposes that verbs in English have an argument structure and are theta-markers. Eventive nominals in English have an argument structure as well, but they cannot assign theta-roles. Therefore, in order to have grammatical arguments, eventive nouns need a theta-assigner. Thus, eventive nouns appear with a preposition in English (e.g. *The donation of money to hospitals*) and with a light verb in some constructions of Japanese. Our proposal is that Kadiwéu does not have any lexical category that is able to theta-assign. The elements that we classify as verbs in Kadiwéu are similar to eventive nominals of better known languages: they have an argument structure but they are not able to license grammatical arguments. In order to take grammatical arguments they need the mediation of elements able to assign theta-roles (see Sandalo 1995 for discussion and see Jelinek & Demers 1994 for arguments that transitive predicates are derived in some Salish languages, which are also polysynthetic).

In summary, we propose that these and possibly other facts about the structure of Kadiwéu would cue the child into the fact that the language is polysynthetic. This conclusion should be sufficient to allow the child to deduce that Kadiwéu allows <e> binding into embedded clauses. This account has the advantage





'<e><sub>i/\*j</sub> knows that <e> j loves John<sub>i</sub>.'

In our Object test sentence (8), *the car* is disqualified as a coreferent for main clause subject <e>, and this privilege is passed down to *Paulo*. It is possible that children initially fail to allow this secondary option requiring a stricter form of parallelism. Of course, it could be that this is simply a preference and might also be exhibited to some extent by adults if one were to test them on the same structures.

## 8. References

- Baker, M. (1995). *The Polysynthesis Parameter*. Oxford: Oxford University Press.
- Crain, S. (1991). Language acquisition in the absence of experience. *Behavioural and Brain Sciences*, 14.
- Crain, S. & McKee, C. (1985). Acquisition of structural restriction on anaphora. In S. Berman, J. McDonough (eds), *Proceedings of the 16th North Eastern Linguistics Society Meeting*. Amherst, Mass: Graduate Linguistics Student Association.
- Jelinek, E. (1984). Empty Categories, Case, and Configurationality. *Natural Language and Linguistic Theory*, 2, 39-76.
- Jelinek, E & Demers, R.A. (1994). Predicates and Pronominals in Straits Salish. *Language*, 70, 697-736.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Gibson, E. & Wexler, K. (1994), Triggers. *Linguistic Inquiry* 25. 407-454.
- Gordon, P. (1985). Level ordering in lexical development. *Cognition*, 21, 73-93.
- Grimshaw, J. (1990). *Argument Structure*. Cambridge, MA: MIT Press.
- Rizzi, L. (1986). Null objects in Italian and the theory of pro. *Linguistic Inquiry*, 17, 501-558.
- Rizzi, L. (1996). *The Fine structure of the left periphery*. Genève: Université de Genève ms.
- Sandalo, F. (1995). *A Grammar of Kadiwéu*. Ph.D. dissertation. University of Pittsburgh, PA.
- Sandalo, F. (1997). *D° incorporation, bare NP extraction, and the organization of the CP layer in a polysynthetic language*. Paper presented at the Linguistics Society of America Annual Meeting, January 3-6. Chicago.
- Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior*, 13, 274-81.
- Speas, M. (1990). *Phrase Structure in Natural Language*. Netherlands: Kluwer Academic Publishers.
- Tavakolian, S. (1981). The conjoined-clause analysis of relative clauses. In S. Tavakolian (Ed.). *Language Acquisition and Linguistic Theory*. Cambridge, Mass: MIT Press.

