Agreement in K'iche' (Mayan): Reflections on Microvariation and Acquisi- tion

Marisa Fried
University of Maryland

Paulina Lyskawa
University of Maryland

Rodrigo Ranero
University of Maryland

Follow this and additional works at: https://repository.upenn.edu/pwpl

Agreement in K'iche' (Mayan): Reflections on Microvariation and Acquisi- tion

Abstract

We explore plausible acquisition trajectories that give rise to microvariation in 3PL agreement realization in K'iche' (Mayan). We report elicitation data with two speakers of the language. For one speaker, all inanimate 3PL arguments control agreement optionally. For another speaker, inanimate 3PL arguments base-generated in a specifier position control agreement obligatorily, while those base-generated in complement position control agreement optionally. Taking England (2011)'s corpus study of the language as a starting point, we show how the interplay of universal and non-universal factors during acquisition might give rise to these different grammars.

Agreement in K'iche' (Mayan): Reflections on Microvariation and Acquisition

Marisa Fried, Paulina Lyskawa, and Rodrigo Ranero*

1 The Puzzle

Variability has been reported in the realization of predicate agreement with third person plural (3PL) arguments in several Mayan languages (England 2011 and references therein). The variability in whether 3PL agreement is banned, realized obligatorily, or realized optionally appears to be governed by several factors. These include surface properties like animacy of the agreement controller (England 2011) and the status of the agreement controller as a surface subject or object (Henderson 2009). In addition, less surface-obvious properties have been implicated, such as whether the agreement controller is base-generated in a complement or specifier position (Lyskawa & Ranero 2020, Levin et al. 2020). Below, we illustrate the optional realization of 3PL agreement in K'iche':

(1) Oj   x-**e-/Ø-**qa-q'aluj   e-keb'   che'.
    1PL   COMPL-**B3PL-/Ø-**A1PL-hug PL-two   tree
    'We hugged two trees.'

In this paper, we examine microvariation in the acceptability of 3PL agreement for different constructions in K'iche'. We present data elicited from two speakers who exhibit different patterns in relation to the phenomenon under discussion, even though they are from the same village (El Novillero in Guatemala). We focus only on inanimate agreement controllers and explore a plausible acquisition trajectory for each speaker, resulting in different grammars. Our starting point to delineate these acquisition trajectories is England 2011's corpus study of K'iche', which showed the following distribution of sentences where 3PL agreement controlled by an inanimate would be expected (e.g., out of 32 instances where object agreement is expected, one example shows agreement):

TRANSITIVE SUBJECT	TRANSITIVE OBJECT	INTRANSITIVE SUBJECT	SUBJECT OF NONVERBAL PREDICATE
1/1	1/32	0/6	3/31

Table 1: K'iche corpus data distribution for inanimate 3PL agreement (England 2011: 402).

Based on the above distribution, England concludes that a discourse preference against inanimate agreement is in the process of becoming grammaticalized in K'iche'. The corpus frequency of different constructions in the table, however, varies widely (e.g., only one data point was attested where an inanimate transitive subject would control agreement).[1]

Furthermore, even though England briefly remarks that agreement with inanimates is grammatical (England 2011: 402), no acceptability judgments are presented. We complement the corpus research with targeted elicitation data in order to formulate an answer for an acquisition question: how do acquirers arrive at a generalization that extends to the data which are very rare in the input?

Our elicitation results show that inanimate agreement is not banned. Most importantly, however, data from one speaker exhibit *obligatory* agreement with transitive subjects and positional subjects.

[1]England (2011) does not provide a fine-grained separation of different types of intransitive subjects (e.g. positional vs. unaccusative). As far as we can tell, she places positional subjects within the non-verbal predicate cell, which also includes nominal and adjectival predicates. The distribution of each of these within the category is impossible to determine. What we can conclude, however, is that examples without agreement here are much more common than examples with agreement.

For another speaker, such agreement is optional. Taking England's corpus as representative of the linguistic input available to a speaker, a puzzle arises: how does a learner acquire obligatory or optional agreement in certain syntactic frames despite limited evidence for it in naturally occurring speech? This puzzle exemplifies Plato's Problem and the poverty of stimulus argument (Chomsky 1965 et seq.), questions with a long history in linguistic theorizing that continue to be ripe for discussion. Here we reflect on insight into acquisition trajectories that can be gained from microvariation at the level of idiolects.

2 Proposal

In this section, we propose a plausible answer to the question of how scarce primary linguistic data might give rise to microvariation in the phenomenon of 3PL agreement (to be laid out explicitly in sections 3-4). Our proposal is built on three components: (i) identifying the set of logical possibilities for the acquirer, (ii) general findings about language acquisition, and (iii) language-specific knowledge about K'iche'. Component (i) lays out the landscape of possible steps an acquirer might take in order to reach a target grammar, while (ii) and (iii) constrain this landscape. Therefore, the more work is done on components (ii) and (iii), the more precise the predictions we could make about language acquisition trajectories and thus possible microvariation in agreement patterns.

We assume that agreement is a critical building block of natural language: infants acquiring their language start out with the hypothesis that all arguments must control agreement obligatorily (including inanimate arguments; see e.g. Chomsky 2001). Let us call this Stage 0.

However, upon exposure to data where arguments do not control agreement, acquirers revise their initial hypothesis. Let us call this Stage 1. The exact direction of the hypothesis revision may be speaker- and data-dependent. For example, if an infant acquiring K'iche' is exposed to active, transitive sentences where an inanimate 3PL object is not cross-referenced on the verb with absolutive agreement, she might revise her hypothesis in several ways:[2]

(2) a. ACROSS-THE-BOARD OPTIONALITY: All inanimate arguments control agreement optionally (vs. animate arguments, which control agreement obligatorily).
   b. IT'S ABOUT BASE POSITION: All base-generated complements control agreement optionally (vs. base-generated specifiers, which control agreement obligatorily).
   c. IT'S ABOUT MORPHOLOGICAL MARKING: All absolutive agreement is optional (vs. ergative agreement, which is obligatory).
   d. IT'S ABOUT OBJECTHOOD: Surface objects control agreement optionally (vs. all other arguments, which control agreement obligatorily).
   e. NO AGREEMENT WITH INANIMATES

Upon exposure to further data, the hypothesis may be revised again (let us call this Stage 2). This revision involves formulating a new hypothesis that is still consistent with the earlier data. For example, if an infant at Stage 1 initially entertained hypothesis (2d) and is then exposed to non-agreeing inanimate (also absolutive) subjects of an unaccusative, she might revise the hypothesis to (2a), (2b), (2c) or (2e). In section 3, we show that one speaker arrived at (2b), while in section 4, we show that another speaker arrived at (2a).

A question that arises is why an acquirer would arrive at one or another target grammar (i.e., settle on (2a) vs. (2b)). Recent research has explored whether all possible revisions are equally likely to be entertained. The choice of one revision over another might be guided by several factors that are not mutually exclusive. First, there are innate and universal heuristics an infant brings to the table (Culbertson et al. 2011, Hudson Kam and Newport 2019). For example, the nature of the language learning mechanism might be biased towards regularization, especially if the input is characterized by a high degree of variability. This heuristic might prioritize the across-the-board types of hypotheses, especially the categorical ones like (2a) and (2e). Second, there are language-external

---

[2]Note that these hypotheses place the microvariation in realization of agreement in different modules of the grammar. For example, (2b) makes reference to syntactic notions like specifier and complement, while (2c) is stated in morphological terms. See Levin et al. (2020) for some discussion.

differences among speakers that influence how they process language (Yu et al. 2013). This individual variation is linked to differences in cognitive processing and personality traits, among other factors. This might give rise to prioritizing different aspects of the data, (2b-d). Third, speakers are highly sensitive to statistical probabilities in the input (Yang 2016). Even a small difference in the input might lead to divergent rule construction or no rules at all (maintaining instead a list of memorized "exceptions"). Finally, there are some patterns in a language for which the input is extremely rare and consistent with multiple competing hypotheses. This situation is proposed to give rise to essentially a dice-roll that determines the ultimate rule (Han et al. 2016).

Taken to its logical conclusion, universal biases and heuristics might suggest that microvariation arising from the same input is impossible. Our work suggests that research on microvariation in possible learner grammars can be informative for this debate. The field of language acquisition usually abstracts away from microvariation. It is rare to have the opportunity to approximate the input, while also describing minutely different grammars between speakers. Given that we do find microvariation, there must be an interplay between universal and non-universal factors (speaker-specific choices like the ones we outlined). With this in mind, we now turn to specific data.

3 Novel Data from Consultant A

Focusing only on inanimate arguments, data elicited from the first consultant show that 3PL objects in a transitive frame control (absolute) agreement optionally:[3]

(3) Oj   x-**e-/ø-**qa-q'aluj   e-keb'   che'.
    1PL   COMPL-**B3PL-/ø-**A1PL-hug   PL-two   tree
    'We hugged two trees.'

3PL (absolutive) agreement is optional when controlled by the sole argument of an unaccusative:

(4) Keb'   wa'is   x-**e-/ø-**tzaq   pa   le   ulew.
    two   pot   COMPL-**B3PL/ø-**fall   on   the   floor
    'Two pots fell on the floor.'

In contrast, inanimate arguments control agreement obligatorily elsewhere. Subjects of transitives must control (ergative) agreement—a 3SG agreement marker and no agreement are rejected:[4]

(5) *Context*: A storm caused two trees to fall through your door. You report the damage.
    Le   keb'   che' x-ø-**ki-/*u-/*ø-**wulij   le   u-chi'   ja.
    the two   tree COMPL-B3SG-**A3PL-/*A3SG-/*ø-**knock.down   the A3SG-mouth house
    'The two trees knocked down the door.'

Positional predicates are a distinct class of unergative one-place predicates in Mayan that generally have stative-like meanings (Henderson 2019, Levin et al. 2020). Just like transitive subjects, the sole argument of positional predicates must control agreement on the predicate:

(6) **E-/*ø-**tzalan   keb'   ab'aj   cho   le   xan.

---

[3]We follow the Mayanist convention of labelling ergative/genitive agreement as set A and absolutive agreement as set B. Abbreviations are as follows: A-set A (ergative/genitive); B-set B (absolutive); COMPL-completive (perfective); DEM-demonstrative; PL-plural; SG-singular. For detailed descriptions of K'iche' syntax and information structure, see López Ixcoy 1997, Par Sapón & Can Pixabaj 2000, Can Pixabaj & England 2011, Can Pixabaj 2017.

[4]When assessing optionality of 3PL agreement through the absolutive paradigm, a potential confound arises, since it is unclear whether the presence of a null morpheme means that there is no agreement whatsoever, or whether there is 3SG agreement (which is null within that paradigm). Within the ergative paradigm, however, we will see that the optionality of agreement is reflected in whether A3PL or A3SG is realized. We assume, then, that a 3SG morpheme, even if phonologically overt, is the surface reflex of a failed agreement relation in the syntax (Preminger 2014, Levin et al. 2020). In other words, we assume that the null morpheme in examples like (1) and the *u*- morpheme in examples like (7) result from failed syntactic agreement.

B3PL-/*Ø-lean   two   stone   against the   wall
'Two stones lean against the wall.'

The empirical generalization for Consultant A is the following:

TRANSITIVE OBJECT (ABSOLUTIVE)	INTRANSITIVE SUBJECT (ABSOLUTIVE)	TRANSITIVE SUBJECT (ERGATIVE)	POSITIONAL SUBJECT (ABSOLUTIVE)
optional	optional	obligatory	obligatory

Table 2: 3PL inanimate agreement in K'iche'; Consultant A.

The data surveyed so far do not exhaust the space of constructions that could be tested (see Levin et al. 2020, Lyskawa & Ranero 2020 for a template of such an investigation in closely related Tz'utujil). However, this sample of data serves to make our point: the generalization for Consultant A is that base-generated specifiers control agreement obligatorily (transitive subjects, positional subjects), while base-generated complements control agreement optionally (transitive objects, un-accusative subjects). Applying our proposal in Section 2 to the final grammar of Consultant A, we assume first that she was exposed to data along the lines of England 2011's sample. We propose that she filtered such data and arrived at the final Stage hypothesis where 3PL inanimate base-complements control agreement optionally, while all other inanimate arguments control agreement obligatorily, (2b). This is identical to a pattern observed for a Santiago Tz'utujil speaker (Levin et al. 2020). Note, crucially, that no other generalization would suffice—e.g., a surface subject vs. object distinction does not govern the pattern, nor does an ergative vs. absolutive distinction.

4 Novel Data from Consultant B

Data from this consultant show a different pattern from Consultant A. For this speaker, all inanimate arguments control agreement optionally. The data below show that transitive subjects control agreement optionally (7), as do positional arguments (8):

(7) Keb'   poro'n   x-Ø-**ki-/u-/*Ø-**k'eso   ri   tijob'al.
      Two   fire   COMPL-B3SG-**A3PL-/A3SG-/*Ø-**destroy   DEM   school
      'Two fires destroyed the school.'

(8) Keb'   ab'aj   **e-/Ø-**q'eel   cho   le   xan.
      Two   stone   **B3PL-/Ø-**lean   against   the   wall
      'Two stones lean against the wall.'

The pattern, summarized in Table 3 below, is similar to closely related Kaqchikel (as reported by Henderson (2009)):

TRANSITIVE OBJECT (ABSOLUTIVE)	INTRANSITIVE SUBJECT (ABSOLUTIVE)	TRANSITIVE SUBJECT (ERGATIVE)	POSITIONAL SUBJECT (ABSOLUTIVE)
optional	optional	optional	optional

Table 3: 3PL inanimate agreement in K'iche'; Consultant B.

Applying our proposal to the grammar of Consultant B, we assume that she was also exposed to the data distribution in England 2011. However, this consultant filtered the data in a different manner than Consultant A: her Stage 1 hypothesis was that *all* inanimate arguments control agreement optionally, (2a). Therefore, she accepts optional agreement in those frames where the data is scarce in the input (transitive subjects and positional subjects).

5 Discussion and limitations

In light of the poverty of the stimulus, what prevents or gives rise to microvariation across different domains in the grammar? There exists a vast logical space of possible ways of parsing the input, coupled with the interplay of universal heuristics and non-universal factors. We would like to suggest that convergence on a pattern might be as puzzling as microvariation, despite the relative lack of attention given to the latter in relation to acquisition. Bringing microvariation into the picture allows us to approximate a controlled experiment and study the interplay of universal and non-universal factors involved in acquisition. This has an advantage over comparing acquisition trajectories across grammars that diverge across several dimensions (e.g., different languages or very divergent dialects; see Kayne 2008 for a similar discussion related to comparative syntax).

Our work is not done here. We acknowledge that a corpus of child-directed speech from El Novillero K'iche' would be more adequate in representing the input that our consultants received. Further, we predict more patterns to be attested in relation to variable 3PL agreement with inanimates than have been documented. Some predictions are listed in (2) but one can construct other logical possibilities. Nevertheless, we hope that this paper will spark interest in documenting microvariation using a targeted elicitation method.

References

Can Pixabaj, Telma, and Nora C. England. 2011. Nominal topic and focus in K'ichee'. In *Representing language: Essays in Honor of Judith Aissen*, ed. R. Gutiérrez-Bravo, L. Mikkelsen and E. Potsdam, 15–30. Santa Cruz: Linguistics Research Center, UC Santa Cruz.

Can Pixabaj, Telma. 2017. K'iche'. In *The Mayan Languages,* ed. J. Aissen, N. England and R Zavala, 461–499. New York, NY: Routledge.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language,* ed. M. Kenstowicz, 1–50. Cambridge, MA: MIT Press.

Culbertson, Jennifer, Paul Smolensky, and Géraldine Legendre. 2012. Learning biases predict a word order universal. *Cognition* 122:306–329.

England, Nora C. 2011. Plurality agreement in some Eastern Mayan languages. *International Journal of American Linguistics* 77:397–412.

Han, Chung-hye, Julien Musolino, and Jeffrey Lidz. 2016. Endogenous sources of variation in language acquisition. *PNAS* 4:942–947

Henderson, Robert. 2009. A case-agreement split in Kaqchikel. Ms., UC Santa Cruz.

Henderson, Robert. 2019. The roots of measurement. *Glossa: A Journal of General Linguistics* 4.

Hudson Kam, Carla. L., and Elissa L. Newport. 2009. Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology* 59:30–66.

Kayne, Richard S. 2008. Some notes on comparative syntax, with special reference to English and French. In *The Oxford Handbook of Comparative Syntax*, ed. G. Cinque and R. S. Kayne. Oxford: OUP.

Levin, Theodore, Paulina Lyskawa, and Rodrigo Ranero. 2020. Optional agreement in Santiago Tz'utujil is syntactic. To appear in *Proceedings of the 41st Annual Conference of the German Linguistic Society*, ed. A. Pankau and V. Struckmeier. URL https://tinyurl.com/8dyp37am.

López Ixcoy, Candelaria Dominga (Saqijix). 1997. *Ri Ukemiik ri K'ichee' Chii': Gramática K'ichee'*. Guatemala: Cholsamaj.

Lyskawa, Paulina, and Rodrigo Ranero. Optional agreement in Santiago Tz'utujil. Ms, University of Maryland.

Par Sapón, Beatriz, and Telma Can Pixabaj. 2000. *Ujunamaxiik ri K'ichee' Ch'ab'al: Variación Dialectal en K'ichee'*. Guatemala: Cholsamaj.

Preminger, Omer. 2014. *Agreement and its Failures*. Cambridge, MA: MIT Press.

Yang, Charles. 2016. *The Price of Linguistic Productivity: How Children Learn to Break Rules of Language*. Cambridge, MA: MIT Press.

Yu, Alan C. L., Carissa Abrego-Collier, and Morgan Sonderegger. 2013. Phonetic imitation from an individual-difference perspective: subjective attitude, personality and "autistic" traits. *PLoS ONE* 8.

Department of Linguistics
University of Maryland
College Park, MD 20742–7502
mfried96@terpmail.umd.edu
lyskawa@umd.edu
rranero@terpmail.umd.edu