

Abbreviations in glosses:

1.2S	portmanteau agreement prefix: 1st person subject, 2sg object
3.	3rd person agreement marked through stem modification
A1S, A2S, A2P	person agreement prefixes for the subjects of transitive and some intransitive verbs
B1S, B2S, A2P	person agreement prefixes for transitive objects and some intransitive subjects
A3, B3	3rd person agreement prefixes (unmarked for number)
A1P.E, B1P.E	agreement for 1pl exclusive
A1P.I, B1P.I	agreement for 1pl inclusive
CNTR	contrastive focus particle
COND	conditional suffix ('if', 'when')
COP	copula (optional)
EVID	unmarked evidential particle
FUT	future suffix (on verbs)
HABIT	habitual
LA	Spanish-derived <i>la</i> – the thing we're interested in
NEG	negative prefix or suffix
NEG.FUT	negative future suffix
N.FUT	future suffix (on nouns)
Q	question marker
PASS	passive prefix
PERF	perfective suffix ('already')
PL	optional plural suffix on nouns
PROG	progressive particle
PURP	purposive suffix ('in order to')
Q.DIST.PAST	distant past particle in questions, also for past events the speaker infers but didn't personally witness
REFL	reflexive prefix
REL	relative clause suffix (on the verb)
SBRD	subordinate clause suffix (on the verb)
SER	serial verb suffix (on the secondary verb in some verb+verb constructions)
TOT	totalitive suffix ('all', 'completely', 'finished')
UNCERT	uncertain future suffix
UNREAL	unreal or counterfactual

Thanks to the speakers I've worked with:

Vicente Cardozo
Luz-María Ojeda
David Barrios
Liza Amarilla
Jazmin Pinazzo
Melki Melgarejo
Evelia Careaga

The speaker for each sentence is indicated in the right margin.

If the ID code for the sentence begins with T, it's from a natural conversation or narrative, or (rarely) a written text or movie.

Guaraní *la* – definitely not a definite article

Kevin Russell, University of Manitoba

Kevin.Russell@umanitoba.ca

PWoLL, Saskatoon, March 18, 2017

Paraguayan Guaraní (Tupí-Guaraní subfamily of the Tupian family) is an official language of Paraguay spoken by about 6 million people. Some factoids:

- very free “word order” for major constituents, though most adverbs and grammatical particles are much fussier.
- “polysynthetic”: often very long verbal complexes, not-quite-fossilized noun incorporation, NPs usually optional.
- kinda head-marking: nouns not marked for core cases (except human patients), but incomplete agreement on verbs. Many verbs select postpositional phrases as arguments.
- split-intransitivity – “subjects” of some intransitive verbs/adjectives have the same agreement as patients of transitive verbs.
- no number distinction in third person verb agreement, but optional plural marking on NPs and number marking on demonstratives.

A majority of speakers are bilingual in Spanish. Code-switching and borrowing (in both directions) are common.

By far the most frequent word originally borrowed from Spanish is *la* ‘the’.

- (1) La óga o- kai -pa -ite. 1b2tn022
VC
LA house A3-burn-TOT-very
‘The house burned completely.’

La is never modified for gender or number, even when used with Spanish-derived nouns, like *típo* < *tipo* ‘guy’ and *vícho* < *bichos* ‘animals’.

- (2) Adio che eíra he’i la típo la kuñá =me. T36a.001
bye my honey 3.say LA guy LA woman=to
“Bye, honey,” said a guy to a woman.’

- (3) Ha ymá -va ningo heta -ve la vicho -kuéra. T34a.113
SV
and long.ago-HABIT EVID 3.many-more LA animal-PL
‘And back then there were more wild animals.’

Spanish articles other than *la* are used only in clear code-switches and in the single idiomatic phrase *lo mitã* (*los* child) ‘the guys, people in general’.

- (4) O- kañy lo mitã =gui. T39e.019
MM
A3-hidden the.PL child=from
‘He was hiding from people.’

1 Guaraní demonstratives

<i>determiner</i>	<i>pronoun</i>	
pe	pé(v)a	‘that, the’, unmarked (or middle-distance)
umi	umí(v)a	‘those, the’, unmarked or middle-distance
ko	kó(v)a	‘this’
ko’ã	ko’ã(v)a	‘these’
upe	upé(v)a	‘that’ (middle-distance)
amo	amó(v)a	‘that, those’ (far-distance)
ku	—	‘that, those’ (absent, or previously mentioned)
aipo	aipó(v)a	‘that’ (unseen but audible, or previously mentioned)

A somewhat common construction uses a demonstrative pronoun followed by the corresponding determiner and the rest of the noun phrase, e.g., *péa pe...*, *kóa ko...*:

- (5) **Kóa** **ko** tape ningo Industrial ára =pe guare. T34a.084
SV
 this.one this road EVID Industrial day=to when
 ‘This road is from the time of Industrial.’

La never does this: **láva la...*

La may come at the beginning of a nominal phrase, before any native demonstratives:

- (6) **La** **pe** mundo o- iko, nde-gustá =rõ térã 5a1kro36
VC
 LA that world A3-exist B2S-pleasing=COND or
 na- nde-gustá -i =rõ.
 NEG-B2S-pleasing-NEG=COND

‘The world exists, whether you like it or not.’

- (7) ... ha che upérõ sa’i a- kopi kuri **la** **ko** To6.068
VC
 and I then little A1S-understand PAST LA this
 Ingle- ñe’ẽ =pe.
 English-language=to

‘... and I only understood a little English back then.’

- (8) Ai- kuaa che **la** **umi** hénte pueblero -kuéra, ... T11.027
VC
 A1S-know I LA those people town.people-PL

‘I know these village people, ...’

2 *La* doesn’t mark definiteness

La is often used with indefinite and/or non-specific noun phrases — as found by SaeMi Choi in our 2012–3 field methods course on Guaraní. In elicitation, speakers often *prefer* the indefinite and non-specific readings, especially if discussing the semantic difference between *la* and a native determiner like *pe* ‘that’

Generic noun phrases are usually preceded by *la*. Native demonstratives can’t occur on a generic noun phrase.

- (9) **La** jaguarete i- ñarõ -iterei. 3b1kro85
VC
 LA jaguar B3-aggressive-very
 ‘Jaguars are vicious.’

(10) Pe jaguarete i- ñarõ -iterei. 3b1kro86
 that jaguar B3-aggressive-very VC

‘That jaguar is vicious.’ (not ‘Jaguars are vicious.’)

(11) Nda-i- katú -i ja- je- rovia la polítiko =re. 3b1kro89
 NEG-B3-possible-NEG A1P.I-REFL-believe LA politician=at VC

‘You can’t trust a politician.’

In negative existential constructions:

(12) Ndaipóri la tembi’u. 270kr111
 there.isn’t LA food VC

‘There isn’t any food.’

(13) * Ndaipóri ko tembi’u. 270kr112
 there.isn’t this food VC

‘This food is all gone.’

La often occurs before *peteĩ* ‘one’ with referents that the speaker doesn’t believe will be identifiable by the listener.

(14) Tape-hũ =gui o- pyta oiméne la peteĩ kilómetro. T07.013
 road-black=from A3-located maybe LA one kilometre. DB

‘From the highway it’s about one kilometre.’

(15) Ha o- hechá=ngo la peteĩ provléma o- reko-ha la T39b.004
 and A3-see =EVID LA one problem A3-have-SBRD LA EC

hapykue goty ládo.
 3.behind toward side

‘He [the doctor] saw that he [speaker’s son] has a problem in his back.’

even non-specific and non-existent referents, like ‘(the) my future house’ in:

(16) A- jogua-se la peteĩ che roga -rã, ha 554kro23
 A1S-buy -want LA one my house-N.FUT and VC

o- je- jogua-pá -ma.
 A3-PASS-buy -TOT-PERF

‘I want to buy a house, but they’ve all been sold already.’

or at the beginning of relative clauses, even those modifying non-specific NPs:

(17) Nde-re- juhu-mo’ái peteĩ óga la nde-gustá -va. 565kro83
 NEG-A2S-find -NEG.FUT one house the B2S-pleasing-REL MM

‘You won’t find a house that you like.’

3 *La* before adjunct clauses

While almost never showing up here in elicitation, in natural texts *la* is very common at the beginning of a conditional clause marked by *-rõ* or *-ramo* on the verb (the same construction can be used for simple temporal cooccurrence):

- (18) Ha la i- ñaña-rõ la o- heká -a hína pláta T09.054
 and LA B3-bad -COND LA A3-look.for-REL PROG money VC
 yvy- guy, nd- o- topá-i.
 earth-under NEG-A3-find -NEG

‘And if the ones looking for the buried treasure are bad, they don’t find it.’

- (19) Ha upé-icha ore ro- japo, la ro- ’ú -rõ la T29.318
 and that-like we.EXCL A1P.E-do LA A1P.E-drink-COND LA DB
 ka’ay ro- ñe’ẽ, ro- kasea heta mba’e.
 mate.tea A1P.E-speak A1P.E-chat a.lot thing

‘And that’s what we do, when we’re drinking mate we talk, we chat about lots of things.’

- (20) ... péro la rei- puru ramo la pakova rogue o- me’ẽ chupe T30.057
 but LA A2S-use COND LA banana leaf A3-give to.3 LA
 peteĩ savor especial.
 one flavour special

‘... but if you use them, the banana leaves give it a special flavour.’

From an autobiographical text in the grammar of Gregores and Suarez (1967: 210–214, orthography standardized), probably narrated by Miguel Azuaga, who was 21 in 1959.

- (21) “Ha guéno, *mi hijo*,” he’i chéve, “ché=ko ro- gratifika T24.008
 and good my son 3.say to.me I =EVID 1.2s-reward
 porã varã la nd- o- jahoga-mo’ãi ramo.”
 good should LA NEG-A3-drown-NEG.FUT COND

‘“Oh, okay, son,” he said to me, “certainly I will have to reward you well if they don’t get drowned.”’

Before other adjunct clauses of time:

- (22) La ja- karu rire, ja- há-ta. 554kr110
 LA A1P.I-eat after A1P.I-go-FUT VC

‘After we eat, we’ll go.’

- (23) La tuicha rire mba’e riré =ko i- jetu’u -ve mo’ã T39b.018
 LA large after what after=EVID B3-difficult-more UNREAL MM
 hína kuri.
 PROG PAST

‘After it got bigger, it would have been more difficult.’

- (24) Upéa o- jehu raka’e la a- nasẽ mboyve. 564kr035
 that.one A3-happen Q.DIST.PAST LA A1S-born before JP

‘That happened before I was born.’

Before a (pseudo-relativized) adjunct clause of place, marked as the focus of a yes/no question:

- (25) Ha la ha'e o- mba'apo-há =pe pio ja T39c.009
MM
 and LA she A3-work -SBRD=to Q already
 o- konsegí=ma la i- permíso o- u hağua.
 A3-get =PERF LA B3-permission A3-come PURP
 'Did she already get permission to come from where she works?'

4 *La* in the middle of main clauses

A sentence spontaneously volunteered by Vicente at the very end of the first session we tried to figure out why *la* was showing up on 'if'-clauses in his and David's conversations:

- (26) A- há-ta la a- jogua. 23nkro21
VC
 A1S-go-FUT LA A1S-buy
 'I'll go shopping.'

In the following sentence, *hasy* is a verb inflected for a third-person subject; it can't possibly act as a noun.

- (27) Péa la hasy. 564kro01
JP
 that.one LA 3.sick
 'That person is sick.'

Two almost adjacent sentences from the Guaraní/Spanish movie, *7 cajas*. Guaraní proper names are *not* normally preceded by *la*:

- (28) Máa=pio la Liz. T37z.142
 who=Q LA Liz
 'Who's Liz?'

- (29) Péa pe ñorsa pio la Liz? T37z.144
 that.one that girl Q LA Liz
 'That girl is Liz?'

This kind of *la* can't come initially in a sentence.

- (30) O- hó-ta la o- karú-vo. 564kro24
JP
 A3-go-FUT LA A3-eat -SER
 'He will go to eat.'

- (31) * La o- hó-ta o- karú-vo. 564kro25
JP
 LA A3-go-FUT A3-eat -SER
 'He will go to eat.'

4.1 Structural *la* and second-position particles

(32) Some “second-position” particles/clitics:

pa	unmarked question
piko, iko, pio	politer question
tiko, tio	reminder question
pipo	exclamative/rhetorical question
niko, ko	(unmarked?) evidential (also <i>ningo, ngo, nio</i>)
ndaje, je	reportative evidential
nipo, nimbo	dubitative evidential
ku	another evidential
nga’u	counterfactual wish
mo	speculative modal/evidential
katu	contrastive focus
voi	a mild emphatic (often non-second)

A majority of sentences in natural discourse have at least one of these, though completely unmarked sentences are also common.

It’s possible for these particles to be stacked, as contrastive *katu* and interrogative *pa* are in:

- (33) — Mba’ère katú=pa ne- rasẽ upéicharõ che rajy. T18.017
 — why CNTR=Q B2S-cry then my daughter
 “‘Then *why* are you crying, my daughter?’”

The most frequent position for this kind of *la* in natural discourse is immediately following a second-position particle.

- (34) Tuichá=pa la o- ky. 565kr088
 big =Q LA A3-rain EC
 ‘Is it raining really heavily?’

But *la* can also occur after something that simply *could* have been fronted before a second-position particle, like *upéi* ‘and then next’ in:

- (35) A- karu, upéi la a- mba’apo. 55qkr101
 A1S-eat then LA A1S-work VC
 ‘I ate, then I worked.’

(36) The closest thing I have to a hypothesis:

- Sentences can be structurally (and intonationally) bifurcated into a focused initial element and a tail/remainder.
- Second-position particles usually (but not always) mark the end of the first element in such a bifurcated sentence.
- *La* may optionally mark the beginning of the tail/remainder.

Some degree of focus seems necessary. A simple unstressed subject pronoun isn’t a good enough sentence-initial host:

- (37) * Che la i- katu a- yta. 565kr106
 I LA B3-possible A1S-swim EC
 ‘I can swim.’

But speakers will often accept putting *la* later in the sentence than after the first element, e.g., after both subject and main verb:¹

- (38) Che ai- katu la a- yta. 565kr115
 I A1s-possible LA A1s-swim MM
 ‘I can swim.’

Speakers will sometimes *insist* on *la* going later. For example, in the following, the progressive marker *hína* is an independent sentence constituent, not another second-position particle stacked onto the second-position evidential *nio*.

- (39) * Ko’áãa nio la hína a- mba’apo. 565kr134
 now EVID LA PROG A1s-work EC
 ‘I’m working now.’

- (40) Ko’áãa nio hína la a- mba’apo. 565kr136
 now EVID PROG LA A1s-work EC
 ‘I’m working now.’

I still can’t explain all cases where speakers reject the earliest possible placement. But perhaps the *option* of a later position suggests multiple bifurcations in the sentence, similar to “recursive formation of marked expressions” discussed by Dooley (1982) for the closely related Mbya Guaraní language of Brazil.²

If a vocative phrase interrupts the sentence after a second-position particle, it feels more natural (or less unnatural) for an additional *la* to group with the second part of the sentence than with the first.

- (41) ?? Mba’é=pa, Peru, la re- japo. 73gkr003
 what =Q Pedro LA A2s-do VC
 ‘What are you doing, Pedro?’

- (42) * Mba’é=pa la, Peru, re- japo. 73gkr004
 what =Q LA Pedro A2s-do VC
 ‘What are you doing, Pedro?’

5 Native determiners acting weird like *la*

It’s possible that some of the stranger jobs performed by *la* used to be performed by native Guaraní determiners. Some speakers will judge as (almost) grammatical sentences where *la* has been replaced by *pe* ‘that’.

- (43) Pe o- ký -ramo, a- pytá-ne che róga =pe. 46qkra156
 that A3-rain-COND A3-stay -UNCERT my house=to MM
 ‘If it rains, I’ll stay home.’

¹But the verb *katu* ‘possible’ is weird in a few other ways that you can ask me about much, much later.

²Paraguayan Guaraní *la* may even be related to (cognate to or calqued on a cognate of) the Mbya marker *ma*, which Dooley (1982) says “has no lexical content, and can be considered simply as an indicator of pragmatic boundaries”. But of the few examples of this *ma* I’ve seen in Dooley’s publications, some are clearly cognate to the Paraguayan Guaraní perfective suffix *-ma* and the distribution of the rest doesn’t seem obviously similar to that of Paraguayan *la*.

- (44) Oi-pota la a- sē. 565kr027
 A3-want LA A1S-go.out MM
 ‘She wants me to leave.’
- (45) Oi-pota pe a- sē. 565kr147
 A3-want that A1S-go.out MM
 ‘She wants me to leave.’

Ku is a strange, infrequent determiner that, like *la*, never varies for number and has no pronominal form (**kúva*).

Ku is a strange, very infrequent, old-fashioned second-position modal/evidential.

6 Conclusion

Paraguayan Guaraní, despite not originally having an /l/ phoneme, borrowed the definite article *la* from Spanish, but

- refused to integrate it into its native system of determiners.
- refused to borrow a plural form.
- refuses to use it to mark definiteness.
- uses it before noun phrases in ways that are impossible in Spanish and probably weren't calqued on any existing Guaraní morpheme.
- uses it to mark sentence and information structure, even when there's no noun in sight, in ways that may not have been calqued on any existing Guaraní morpheme.
- *never* uses it obligatorily.
- speakers still consciously believe *la* is “Spanish”, despite its frequency of use having no relationship to a speaker's bilingualism or ethnicity.

A single sentence from a natural conversation, illustrating each kind of *la* we've discussed:

- (46) Ha o- japo jey arã chupe hína la estudio-kuéra T39b.010
 and A3-do again should to.him PROG LA study -PL EC
la o- hecha haḡua, la péa =pa ha'e la provléma o
 LA A3-see PURP LA that.one=Q COP LA problem or
 mba'é=pa la o- je- japo arã.
 what =Q LA A3-PASS-do should

‘He's going to have to do more tests on him in order to see if that's the problem or what can be done.’

- la estudio-kuéra* basic determiner on a noun phrase
 (a borrowing that would have been masculine and plural in Spanish)
la ohecha haḡua before an adjunct clause of purpose
la péa on a noun phrase consisting of a native demonstrative pronoun
la provléma basic determiner
mba'épa la structural, after a second-position particle

And a completely artificial, but grammatical, sentence:

La olala la olala laláa la olala lala.

‘The complainer who's always complaining keeps on complaining.’

References

Dooley, Robert. 1982. Options in the pragmatic ordering of Guaraní sentences. *Language* 58, 2, 307-331.

Gregores, Emma, and Jorge A Suárez. 1967. *A description of colloquial Guaraní*. The Hague: Mouton.