2025. Proc Ling Soc Amer 10(1). 5899. https://doi.org/10.3765/plsa.v10i1.5899.

Distribution of Evidential Markers in a Cuzco Quechua Corpus

Jefferson Imbaquingo-Ramón, Liliana Sánchez, Daniel Perez, & Elena Koulidobrova[1*]

Abstract. Due to contact, several varieties of Quechua are losing evidential markers. Using a Cuzco Quechua corpus (Macedo et al., 2022), we present the distribution of evidential markers *-mi* (*-n*, *-ni*), *-si* (*-s*, *-sis*), and *-chá* on discourse connectors (DCs) (Sanchez et. al., 2021) and sentence-level constituents (subject, object, verb, and adverb). To determine how DCs and constituents are marked with evidentiality, we analyze a corpus of semi-structured interviews in Quechua (N=29; 995 utterances total). 3.42% of evidential suffixes were found on DCs and 96.58% on constituents indicating that their syncretic nature (focus and evidentiality) is at play in constituents. A GLMM suggests that evidential type (*-mi*, *-si* and *-chá*) is predictors of DC and constituent marking. *-mi* is more likely to appear in both syntactic positions, while *-si* shows the greater shift from DC to constituents ($p<0.05$). Men are also more likely to mark DCs and less likely to mark constituents.

Keywords. Evidentiality; focus; discourse connectors; Cuzco Quechua; speaker gender, constituent marking.

1. Evidentiality in Cuzco Quechua. According to Aikhenvald (2012), evidentiality corresponds to the grammatical marking of how knowledge is acquired by a speaker (through visual evidence, evidence from other senses or through inference based on logical assumptions (Aikhenvald 2012: 248). A crucial aspect of Aikhenvald's view of evidentiality is her assumption that a clause lacking an evidential marker would sound strange to a native speaker of Quechua (Aikhenvald 2012: 248).

Importantly, evidential markers can be syncretic with focus (Muysken 1995, Sánchez 2010), namely that their interpretation may involve evidential values as well as focus. This fact holds for three types of sentence-level evidential suffixes in Cuzco Quechua (Cusihuamán 2001), also found in other Quechua languages of the QII family (Parker 1963, Torero 1964, Cerrón-Palomino 1989): attested (*-mi*, *~-m*,*~-n*,*~-mim*, (1)), reportative (*-si-*, *~-s*, *~-sis*, (2)), and dubitative–based on insufficient evidence (*-chá*, *~-ch*, (3))[2]:

(1)	Coronayuq-mi	ka-shka-n	ama
	Corona-POSS-ATT.EVID;FOC	have-PROG-3.SG	NEG.JUSS

[1*] We thank all participants in the study as well as Bersi Macedo, for her work as a community researcher for conducting interviews and transcribing and translating into Spanish parts of the corpus. Thanks also to community researchers César de la Cruz and Braulio Quispe who worked on transcriptions and translations into Spanish and in the case of Braulio Quispe we thank him for coding evidential markers. This research was supported by the National Science Foundation under Grants 2033712 (Sánchez) and 2033739 (Koulidobrova). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Authors: Jefferson Imbaquingo-Ramón, University of Illinois Chicago (jimbaq2@uic.edu), Daniel Perez, University of Illinois Chicago (eperez92@uic.edu), Liliana Sánchez, University of Illinois Chicago (lesanche@uic.edu) & Elena Koulidobrova, Central Connecticut State University (elena.koulidobrova@ccsu.edu).

[2] Adapted from on Leipzig (2015) glossing system: ATT.EVID: attested evidential
REP.EVID: reportative evidential
DUB.EVID: dubitative evidential

achhuy-ku-nkichik-chu.
get.close-REFL-PRES.2PL-NEG
'To those who have Coronavirus, do not get close.'

(2) Corona-yuq-si …
Corona-POSS-REP.EVID;FOC …
'To those who allegedly have Coronavirus, do not get close.'(hearsay)

(3) Corona-yuq-chá ...
Corona-POS-DUB.EVID;FOC …
'To those who (maybe) have Coronavirus, (maybe) do not get close.'

In cases like (1), the attested evidential marker *-mi* (and its allomorphs *-m,-n,* and, *-mim*) and the focus-fonted constituent on which the marker appears as the last suffix, has both an evidential and a focus interpretation. The same holds for the reportative evidential *-si* and its allomorphs, and the evidential/conjecture marker *-chá* and its allomorphs.[3]

In addition to marking constituents, evidential suffixes in Cuzco Quechua may also appear on Discourse Connectors (DC), and as in the case of constituents, they appear in word final position at the right edge. Some DCs have as roots demonstratives (e.g., *chay* "that") and adverbs (e.g., *hina* "like","so") followed by switch reference markers and/or evidential markers (Muysken 1995, 2015; Andrade et al. 2020), as shown in (4)-(6) for each one of the evidential markers in (1)-(3), respectively:

(4) Hina-spa-*n* concienciya-ta hap'ina-nku-paq.' *-mi*
Thus-SS-ATT.EVID;FOC consciousness-ACC take-3;PL-BEN
'Thus (based on previous *p*) so that they change their minds.'

(5) Hina-spa-n-*si* concienciya-ta… *-si*
Thus-SS-3;S-REP.EVID;FOC consciousness-ACC
'Allegedly (based on previous *p*) so that they change their minds.'

(6) Hinaspa-*chá* concienciya-ta… *-chá*
Thus-DUB.EVID;FOC consciousness-ACC
'Maybe (based on previous *p*) so that they change their minds.'

In some cases, an evidential marker can appear on a DC and on a constituent and, while they usually agree in evidentiality values, depending on the context they may not exhibit the same evidential value, as shown in (7):

(7) Chaypi-*s* chiqaq-*mi* hap'i-ru-n.
Thus-REP.EVID true-ATT.EVID;FOC catch-PERF-3S
'Thus (allegedly) they caught (it) in reality.' (Macedo et al. 2022)

Finally, *in-situ* constituents can host evidential markers:

(8) Pirdu wasi-ta-*n* ruwa-n.
Pirdu house:ACC-FOC;EVID build-3S
'It is the house that Pirdu builds.' (attested information) (Muysken 1995)

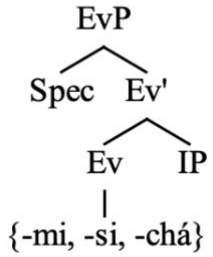
[3] While these evidential markers can still be found in the Southern varieties of Quechua spoken in Peru, they have become less frequent in some varieties of Bolivian Quechua (Kalt 2021) and Ecuadorian Kichwa.

In this paper we demonstrate that in oral production, evidentials may appear on either discourse connectors (DCs) or on sentence-level constituents; our data reveal that the variation in the frequency with which evidential markers appear in discourse depends on their location – on DCs or constituents. Some of this variation is predicted by gender.

2. Assumptions: Evidentials as enclitics on left-most constituents and DCs.

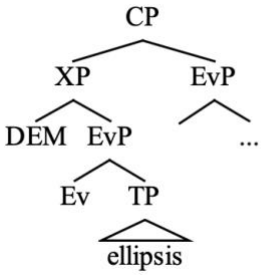
In this paper we make several assumptions. First, we assume Muysken (1995) proposal for evidential markers as enclitics on the left-most constituent in the sentence (9).

(9) Enclitics on the left-most constituent (Muysken 1995)



This analysis, for instance, accounts for the fact that DCs scope over the whole sentence – i.e. the whole sentence conveys the speaker's stance with respect to attested or hearsay information or the conjectural nature of the utterance (as in (4)-(6)). However, in order to account for the distribution of evidential markers in cases like (7), a different syntactic representation is needed – one in which the DC can be independent from the constituent that is marked with an evidential suffix. One such option is (10) (Sánchez, in prep.), – a representation in which the DC, as a discourse anaphora, occupies a hanging topic position that allows for the projection of another Evidential Phrase different from that of the main clause (Villa-Garcia 2023, et seq). The main idea behind this representation is to account for the fact that a DC is marked with evidentiality and has scope over an elided sentence that summarizes what has been said before and is therefore interpreted as a discourse anaphora (Muysken 2015).

(10) Discourse Connector



(Sánchez, in prep)

Second, in this paper, we adopt Sánchez' (2010) analysis of evidential markers as the spell-out of long distance Agree relationships. Evidence in favor of this analysis comes from *in situ* constituents, as shown in (11):

(11) Pirdu wasi-ta-*n* ruwa-n.
Pirdu house-ACC-FOC;EVID build-3S
'It is the house that Pirdu builds.' (attested information) (Muysken 1995)

According to Sanchez (2010), *in situ* marking is possible in languages like *CUZ* both for evidentiality and focus because of morphological syncretism which allows the spell out of the proposed long-distance Agree relationships (Baker 2008) that satisfy both the evidential and the focus values required in the sentence, as illustrated in (12):

(12) [$_{EvidP}$ Evidential [$_{FocP}$ Focus…[vP XP $_{(evid, foc)}$]]] (Sánchez 2010)

The enclitic analysis applies when evidentiality is the only feature valued, as in the case of DCs, or when focus, as is the case with *wh*-words, is the only feature valued. Focus on *wh*-words is not only morphologically marked; it also requires leftward movement as shown in (13a-b).

(13) a. Ima-ta-*m* muna-nki
What-ACC-FOC want-2P
'What do you want?'
b. [$_{FocusP}$ Imata–m [$_{WhP}$ ~~imata~~ [vP…~~imata~~]]]

Furthermore, in terms of its semantic interpretation, we adopt Sánchez et al.'s (2021) proposal, according to which the encoding of evidentiality on the DCs is a way for the speaker to establish their contribution to the Common Ground (Grzech 2020) and their commitment to the ground assigned to the first sequence of events in a narrative or to renew it in subsequent discourse units (Sánchez et al. 2021).

Finally, we follow Harvey (2020), who analyzes discourse practices of men and women in Ocongate (Southern Peru). The author argues that among bilingual speakers among bilingual speakers of Cuzco Quechua, differences in the speech of gender-based groups lie in the "latent meanings" (Harvey 2020) and affect interpretation of discourse proper.

3. Study: Evidentials in the SAIPM-COVID19 corpus.

3.1. Research Questions. As noted by Aikhenvald (2021) the preferences that govern the choice for evidential markers are linked with access to knowledge and, she argues, can be manipulated in discourse to reflect the speaker's stance or attitudes. Further, how speakers come to acquire or express knowledge, or their level of expertise, may be modulated by some social factors (Grzech 2020 and Aikhenvald 2021). In this study, we explore two factors that may influence speaker choice to establish or renew their commitment to the grounds for knowledge or information: i) contact with a socially dominant language that lacks evidential markers (such as Spanish) and ii) the gender. The study is guided by the research questions in (14):

(14) RQs
1. How frequently do evidential markers appear on discourse connectors and on constituents?
2. Which evidential marker(s) (*-mi*, *-si*, *-chá*) are more likely to appear on DCs?

3. Does gender of the speaker and contact with Spanish have an effect on the syntactic position of the evidential marker?

Regarding RQ1, we hypothesize that evidential markers will appear more frequently on constituents as they tend to be syncretic with focus to mark new information increasing the need for marking. In DCs, only evidentiality will be marked as DCs are only required to establish or renew the speaker's commitment to the grounds for assertion at the discourse level.

Regarding RQ2, we predict that *-mi* should be the most frequent evidential marker. This is because *-mi* marks both evidentiality and grammatical focus, thus increasing the potential frequency of tokens in the dataset. Additionally, the stimuli (see section **3.2)** principally elicit a ranked order of evidentials: *attested* (= "what I know to be true") > *reportative* ("what what people say") > *conjectural/dubitative* ("what I am not sure about").

Finally, regarding the RQ3, we hypothesize that gender will be a predictor of how frequently the speaker establishes or renews their commitment to the grounds for the information expressed. More generally, gender may have an effect on distribution of evidential markers.

3.2. Methodology. This study is part of a broader project: Salud Indígena, Poblaciones Minorizadas, y COVID19 (SAIPM-COVID19) (*Indigenous Health, Minoritized Communities and COVID19*). Data were collected among two linguistic groups in Peru: speakers of Cuzco Quechua from seven rural communities in the Cuzco region (provinces of Acomayo, Anta, Calca, Canchis, Paucartambo, Paruro, and Quispicanchis; N = 71) and speakers of Shipibo-Konibo from two rural communities in the Ucayali region (N = 38). This study focuses on the subset of the Quechua participant pool, highlighted in Figure 1 (N = 29, 41% of the Quechua dataset)..

Figure 1. Quechua communities interviewed

Following the principles of Participatory Action Research (PAR), the research team collaborated with community researchers (members of the indigenous communities and native speakers of the respective languages) to develop a semi-structured interview tool (Sánchez and Koulidobrova, 2022). This questionnaire is publicly available online (https://sites.google.com/view/saipm-covid19/home). The interview covered various topics including demographic and linguistic background, language ideologies, and perspectives on the ongoing crisis. It consisted of open-ended questions and multiple-choice items, aimed at assessing interviewees' responses to statements originally produced by the World Health Organization as warnings against COVID-19. Each interview yielded 30-55 min of audio data which contain

information related to the participants demographics, language experiences and attitudes, perceptions related to the health-crisis and its prevention, and information about channeling preferences and community preparedness. All data were collected by community researchers who conducted interviews in person, via Skype, or by phone, either using google forms or pen and paper (see Figure 2).

Figure 2. Indigenous Language Questionnaire.

The interviews were conducted in the Indigenous language when possible. Indigenous language proficiency was assessed by community researchers – trained intercultural bilingual educators with extensive experience. A new rubric was created for this assessment deliberately (Sánchez and Koulidobrova, 2020); community researchers were consulted in the creation and trained in the application of the rubric (Table 1).

Criteria for assessment: Ability to conduct an interview in [primary indigenous language].

1. This conversation is impossible in [primary indigenous language], and the interview was switched to the [dominant language of the country]	Beginner: low + med = 1
2. This conversation is difficult to have in [primary indigenous language]. The participant understands the language and it is possible that in their answers, the participant switches to [dominant language of the country] a lot.	Beginner: high =2
3. This conversation is possible to have in [primary indigenous language]. However, the participant does not know many words I ask. They can easily talk about things they know about, but when the questions are difficult, or when we talk about Corona, they stumble and switch to the [dominant language of the country].	Intermediate: low + med = 3
4. This conversation is easy to have in [primary indigenous language]. However, for some words/ideas it is possible that the participant will need to switch to [dominant language of the country].	Intermediate high – Advanced low = 4
5.This person can say everything in [primary indigenous language].	Advanced high = 5
	= native/native-like

Table 1. Holistic rubric for Indigenous language assessment

No such rubric was utilized for the dominant language (Spanish); thus, other measures stand proxy for knowledge of/contact with Spanish. For instance, several questions in the interview target contact with Spanish directly – exposure to it during childhood and exposure to it during various times of the day: at work, with friends, family, in community, etc (daily contexts numbering 1-4). For the purposes of this study, the latter was taken as the measure of contact.

3.3. Coding. All the interviews were transcribed, translated into Spanish and English, and are publicly available, both in pdfs, docx, mp3s, and ELAN (Cuzco Quechua corpus, Macedo et al., 2022). To ensure the accuracy and consistency of the data annotation, the coding process was conducted by native speakers of Quechua. Following the initial coding, two principal investigators (PIs), who are L2 learners of Quechua, reviewed the data for verification. Inter-rater reliability was assessed, yielding an agreement rate >85% between the native speaker coders and the PIs. Any discrepancies identified during this process were thoroughly discussed and fully resolved, ensuring 100% agreement on the final dataset.

3.4. Results. The descriptive characteristics of the participant pool are described in Table 2.

Participant	Age	Gender	Number of contexts, Spanish use (0-4)	Number of contexts, Quechua use (0-4)	Quechua proficiency (1-5)
1	51	M	0	4	5
2	39	M	0	4	5
3	56	F	4	4	4
4	74	M	4	4	4
5	61	F	0	4	5
6	47	F	0	4	5
7	19	F	4	1	3
8	26	F	0	4	5
9	30	F	4	4	4
10	44	M	3	4	4
11	38	F	4	4	5
12	25	F	1	4	4
13	37	M	1	4	5
14	42	F	2	4	4

15	63	M	3	2	5
16	52	F	0	4	5
17	59	F	0	4	5
18	40	M	1	4	4
19	56	M	2	4	5
20	51	M	2	4	4
21	53	M	1	4	5
22	77	F	1	4	5
23	63	F	0	4	5
24	37	F	0	4	5
25	37	F	2	4	5
26	63	F	0	4	5
27	55	F	0	4	5
28	18	F	0	4	4
29	60	F	0	4	5

Note: "Language context" stands for the types of discourse situations, e.g. "at work", "with the family," "with the community" and "with friends." The number of contexts ranges from 0 to 4. Proficiency ranges from 1 to 5.

Table 2. Participants' characteristics

The resulting dataset analyzed consists of 1,175 sentences containing sentence-level evidential markers. Examples (14)–(16) illustrate sentences from the dataset that contain evidential markers within constituents. These examples demonstrate how evidential markers function within noun phrases, verb phrases, or other syntactic units.

(14) Hamu-chka-n-*mi* unquy
Come-PROG-3.S-ATT.EVID;FOC illness
'An illness is coming.'

(15) Achwan-*si* wañu-chka-nku
A lot-REP.EVID;FOC die-PROG-3;P
'A lot more people are dying'

(16) Chay-qa chiqaq-puni-*chá*
That-TOPtrue-DEF-DUB.EVID;FOC

chay-ri ka-n-man
that-RESP be-3.S-COND
'That might always be true'

Similarly, (17)–(19) show evidential markers on DCs. These cases illustrate how evidentiality interacts with discourse structure, signaling information source and speaker stance in transitional or organizing elements within speech.

(17) Chay-*mi* wawa-yku-ta tapuri-ku-yku
Thus-ATT.EVID;/FOC child-1;P;EXCL-ACC ask-POL-1;P;EXCL
'Thus we only asked our children.'

(18) Chay-manta-*s* chay paqari-mu-n-man
Thus-ABL-REP.EVID this appear-TLOC-3;S;CON
'Thus, it is said that it would appear (like this).'

(19) Chay-*chá* runa wañu-n achka riki
Thus-DUB.EVID people die-3.SG many right?
'Thus, a lot of people are dying, right?'

Table 3 presents the distribution of evidential markers by syntactic position, revealing the preference for marking constituents over discourse connectors.

Evidential Marker	On constituent	On discourse connector
First-hand (*-mi*)	882	38
Reportative (*-si*)	55	7
Dubitative (*-chá*)	189	4
TOTAL	1126	49

Table 3. Distribution of evidential markers by position

Overall, these frequencies suggest that evidential markers overwhelmingly appear in constituents (n = 1126, 95.8%), with DCs accounting for only 4.2% of the cases. Among the three evidential types, *-mi* is the most frequent in the corpus. However, *-si* shows the highest relative frequency of DC placement (accounting for 11.3% of its occurrences compared to just 4.1% for *-chá* and 4.1% for *-mi*).

3.5. Analysis. These descriptive patterns suggest that evidential type may influence syntactic position, particularly for *-si*. Several analyses were conducted on the data; here we report only a few.

Two chi-square tests of independence were conducted to examine the relationship between the syntactic position (DC vs. InConst) and two factors: evidential type (*-mi*, *-si*, *-chá*) and gender (F vs. M). The first chi-square test assessed whether the distribution of evidential types varied by syntactic position. The results were significant, X^2 (2, N= 1175) = 9.99, p= .007, indicating that evidential type is not equally distributed across syntactic positions. A second chi-square test examined the relationship between gender and syntactic position. This test also yielded a significant result, X^2 (1, N= 1175)= 6.85, p= .009 suggesting that gender influences the distribution

of DCs and in constituent positions. Together, these findings indicate that both evidential type and gender might play a role in determining the syntactic position of evidential markers. However, it remains unclear whether contact with Spanish plays a role in the syntactic position of evidential markers.

To assess the effects of Spanish contact (measured on 1-4 scale), gender , and evidential type on the syntactic position, the data was entered into a Generalized Linear Mixed Model (GLMM) with syntactic position as the outcome variable. The model included S*panish contact, gender* and *evidential type* as fixed effects, along with their interaction. *Participant* was included as a random intercept to account for individual variability. Table 4 shows the summary of the GLMM findings.

Predictor	**Estimate (log odds)**	**SE**	**z-value**	***p*-value**
(Intercept)	-3.372	0.302	-11.17	< .001***
EvType(-*si*)	1.055	0.464	2.274	0.023*
EvType(-*chá*)	-0.451	0.552	-0.817	0.414
SpanishContext	-1.256	0.917	-1.370	0.171
Gender:Male	0.637	0.589	1.082	0.279
SpanCont:Gender	1.112	0.953	1.168	0.243

***Note:** *p < .05; **p < .01; ***p < .001. The intercept represents the expected log-odds when all predictors are at zero.*

Table 4. Fixed effects from the GLMM

The model revealed a significant effect of evidential type, particularly for the evidential *-si*, which was more likely to appear in the discourse connectors position compared to the *-mi* and *-chá* ($\beta = 1.05$, $SE = 0.46$, $z = 2.27$, $p = .023$). Neither Spanish context ($\beta = -0.70$, $SE = 0.48$, $z = -1.47$, $p = .141$) nor gender ($\beta = -0.64$, $SE = 0.59$, $z = -1.08$, $p = .279$) had a significant main effect on the syntactic position of evidential markers. This suggests that the likelihood of placing an evidential marker in a discourse position does not meaningfully differ between male and female speakers, nor does it systematically vary with the number of Spanish contexts a speaker is exposed to (see Table 2). The model's goodness-of-fit was assessed using the Akaike Information Criterion (AIC = 398.8). The model's log-likelihood was -192.4. The model explained 36% of the variance in syntactic position when considering both fixed and random effects ($R^2 = 0.36$). This suggests that while evidential type plays a role in predicting syntactic position, additional individual-level factors may contribute to variation.

Figure 3 below further illustrates the relationship between the predictors and the likelihood of in constituent vs. in discourse connectors placement as it presents the predicted probabilities for each evidential type (*-mi, -si, -chá*) as a function of Spanish contact, and gender. The predicted probability estimates were derived from the fitted GLMM model and demonstrate visually how different predictor combinations influence the syntactic position of evidential markers.

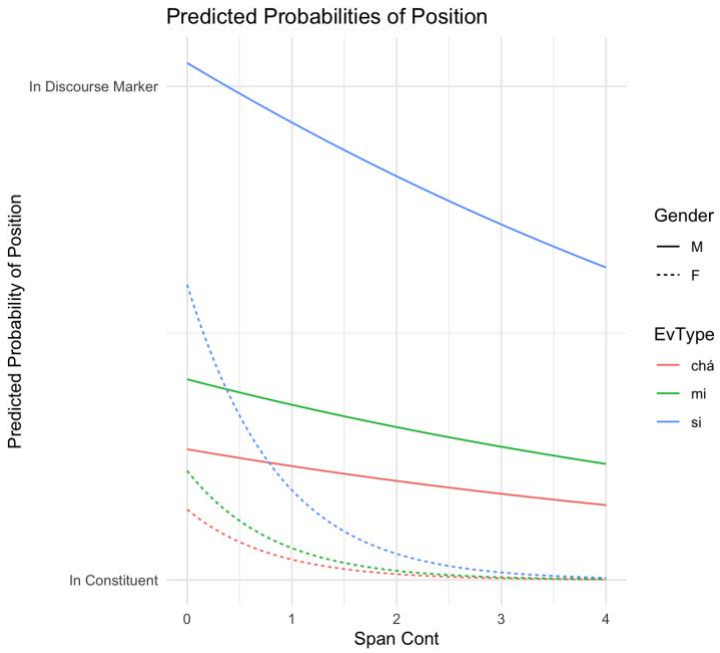


Figure 3. Predicted Probability of Syntactic position

The plot reveals an effect of evidential type on syntactic position. The probability of discourse connector placement is highest for *-si* (blue line), particularly when the Spanish context is low (Spanish contact = 0). As the presence of Spanish increases from 0 to 4, the probability of discourse connector placement decreases, with this effect being steeper for *-si* compared to *-chá* and *-mi*. Notably, the probability of in-discourse connectors placement for *-chá* (red) and *-mi* (green) is consistently lower, with *chá* exhibiting the lowest probability overall. This suggests that greater exposure to Spanish strongly favors in-constituent placement. Additionally, female speakers (dash lines) consistently showed a lower likelihood of using evidential markers in DCs, with the largest difference observer for *-si*.

Discussion and conclusion

Recall that the goal of the study was to examine the influence of Spanish (or, rather, its contact), gender (F vs. M), and evidential type (*-mi*, *-si*, *-chá*) on the syntactic position of the evidential markers (i.e., on DCs vs. on constituents). Specifically, we addressed the following questions:

(20) RQs: = (14)

1. How frequently do evidential markers appear on discourse connectors and on constituents?
2. Which evidential marker(s) (*-mi*, *-si*, *-chá*) are more likely to appear on DCs?
3. Does gender of the speaker and contact with Spanish have an effect on the syntactic position of the evidential marker?

As an answer to RQ1, we had predicted the prevalence of *-mi.* This was at least partially because the stimuli has essentially elicited the distribution which ranks approximately as “what I know to be true” > “what what people say” > “what I am not sure about". In other words, in terms of evidential markers, we expected *attested* > *reportative* > *dubitative/conjectural.* However, the prediction was also based on syntax, not just the frequency analysis. To that end, Table 3 shows

that the evidential markers overwhelmingly appear on constituents (95.8%) while DCs only account for 4.2% of the total evidentials. Constituents were more frequently marked with the first hand evidential marker *-mi,* which follows the rich syncretic nature of such marker (Sánchez, 2010; Muysken 1996). The high frequency of *-mi* in constituents in comparison to the other evidential markers (*-si* and *-chá*) might suggest that speakers are stating information that has been learnt via first-hand experience and that some of that information is the focus of the information/conversation. This reinforces the role of these markers, especially *-mi*, as not only evidential markers, but also focus markers. Further, the results of the analyses have clearly suggested that while evidential type plays a role in predicting syntactic position, additional individual-level factors may contribute to variation.

We were also able to answer RQ2. Among the three evidential markers under examination (*-mi, -si,* and *-chá*), *-si* is the most frequently positioned as a DC and its rate of suppliance is much higher than that of the other two evidential markers (and their allomorphs): *-si* appears on DCs 11.3% of the time, while *-cha* and *-mi* do 4.1% total. This distribution suggests that while evidential markers mainly appear in constituents, their use in discourse connectors varies by type.

In terms of R3, the dataset offers the following:

Overall, establishment and renewal of the commitment to the ground occurs more consistently within contexts where reported information is being transmitted (hence the higher frequency of *-si*). Regarding speakers' gender, men establish/renew the commitment to the grounds more often than women. Additionally, although Spanish contact did not have a significant effect on the syntactic position of evidential markers overall, an observable trend emerges: the frequency of constituent marking decreases with greater contact with Spanish (see Figure 1).

Quechua uses evidential markers (e.g., *-mi*) and word order (e.g., focus *in-situ*) to mark focus. Similar mechanisms have been documented in Spanish in contact with Quechua, where evidentiality is often expressed through specific lexical items (e.g., *pues*), syntactic reordering (e.g., movement to sentence-initial position or in-situ focus), and verbal morphology (e.g., present perfect tense) (Escobar & Roy 2022; Sánchez 2004; Zavala 2001, a.o.). These examples show the Quechua syntactic structures can influence Spanish. However, we also argue that the contact with Spanish can also impact the system of evidentiality marking in constituent and discourse connectors. Focus marking using intonation or syntactic position of constituents in Spanish seems to be the most frequent strategy. Since evidentiality and focus are syncretic in Quechua, their marking may be affected by how focus marking works in Spanish, namely, by occurring mostly at the constituent-level (Colantoni & Sánchez 2021). While some reportative evidential markers have emerged in Spanish in contact with Quechua as phonologically independent words such as *pues* (Zavala 2021) or *dice* (Sánchez 2004), it is still conceivable that they are lower in frequency because they also compete with non-evidential discourse connectors in Spanish. If speakers are exposed to Spanish more often, such strategies will be more readily available to them, influencing evidential marking in Quechua in consequence, as noted by recent accounts of crosslinguistic influence that highlight the role of co-activation in bilingual development (Sánchez 2019).

To summarize, in this exploratory paper we have presented data that indicates that marking of evidentiality is more frequent on constituents than on DCs**.** This may indicate that their syncretic nature (focus and evidentiality) is more at play on constituents than on DCs. The data also shows an increase in constituent marking vs. DCs marking that is consistent with more contact with Spanish, as more opportunities to speak Spanish seem to influence the marking across the board. The reportative marker *-si* seems to shift the most from DC to constituent marking as contexts of interaction in Spanish increase. More research is needed as the low frequency of *-si* may be

skewing the results. This low frequency can be attributed to the nature of the task used for elicitation (a questionnaire that in many cases inquired about personal experiences with COVID-19). Despite the low frequency of *-si*, the data also seems to indicate that the establishment and renewal of the commitment to the grounds occurs more consistently within contexts where reported information is being transmitted. Finally, the data also seems to indicate that men establish/renew the commitment to the ground more often than women.

References

Aikhenvald, A. (2018). Evidentiality: The Framework. In A. Aikhenvald (Ed.), *The Oxford Handbook of Evidentiality* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198759515.013.1

Aikhenvald, A. (2021). *The Web of Knowledge: Evidentiality at the Cross-Roads*. Brill. https://doi.org/10.1163/9789004466425

Cerrón-Palomino, R. (1987). *Lingüística Quechua*. Centro Bartolomé de no Casas.

Colantoni, L., & Sánchez, L. (2021). The Role of Prosody and Morphology in the Mapping of Information Structure onto Syntax. *Languages*, *6*(4), 207. https://doi.org/10.3390/languages6040207

Escobar, A. M., & Roy, J. (2022). Grammaticalization in the Evidential pathway of the Andean Spanish Present Perfect: Language contact as a trigger for language change. *Isogloss*, *8*(4), 1-38.https://doi.org/10.5565/rev/isogloss.184

Harvey, P. (2020). The presence and absence of speech in the communication of gender. In *Bilingual Women* (pp. 44-64). Routledge.

Kalt, Susan E. "Acquisition, Loss and Innovation in Chuquisaca Quechua—What Happened to Evidential Marking?." *Languages* 6, no. 2 (2021): 76. https://doi.org/10.3390/languages6020076

Lee, Seunghun J., Daisuke Shinagawa, Elena Koulidobrova, Liliana Sanchez. 2022. Health Questionnaire Digital Archive (HealthQDA), Digital collection managed by Information Resources Center at Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies. URL: https://healthqda.aa-ken.jp

Macedo, Bersi, Cesar de la Cruz, Braulio Quispe, Liliana Sanchez, and Elena Koulidobrova(2022) Quechua (CUZ) Health Questionnaire Metadata (ICULD-0037), In: Colon-Rodrigues, Keyra and Jeff Imbaquingo (Eds.) International Christian University Working Papers in Linguistics 20: Health Questionnaire Digital Archive. Tokyo, Japan: International Christian University. pp. 11-753. [LINK: http://doi.org/10.34577/00005149]

Muysken, P. (1995). Focus in Quechua. In *Discourse configurational languages*. New York: Oxford University Press.

Muysken, P. (2015). From nominal predicate to deictic clausal highlighter: The developement of hina “like.” In M. Manley & A. Muntendam (Eds.), *Quechua Expressions of Stance and Deixis* (1st ed., Vol. 11, p. 338). BRILL. https://ebookcentral.proquest.com/lib/uic/detail.action?docID=1991829

Parker, S., & Weber, D. (1996). Glottalized and Aspirated Stops in Cuzco Quechua. *International Journal of American Linguistics*, *62*(1), 70–85. https://doi.org/10.1086/466276

Sánchez, Liliana. (2010). The morphology and syntax of topic and focus minimalist inquiries in the Quechua periphery. In *The morphology and syntax of topic and focus minimalist*

inquiries in the Quechua periphery. John Benjamins Pub. Company. https://doi.org/10.1075/la.169

Sánchez, L., Andrade Ciudad, L. Gonzalo Segura, R. (2021). Evidential/focus markers and discourse connectors in Southern Quechua: Effects of age of acquisition, first language status, and patterns of language use. International Pragmatics Association. June 29th.

Torero, A. (1964). Los dialectos quechuas. *Anales Científicos de la Universidad Agraria*, *11*(2:4), 446–478.

Villa-García, J. (2023). Hanging Topic Left Dislocations as extrasentential constituents: toward a paratactic account. Evidence from English and Spanish. The Linguistic Review, 40(2), 265–310.

Zavala, V. (2001). Borrowing evidential functions from Quechua: The role of pues as a discourse marker in Andean Spanish. *Journal of Pragmatics*, *33*(7), 999–1023.