Chapter 4

The realizational coefficient: Devising a method for empirically determining prominent positions in Conchucos Quechua

Timo Buchholz
Freie Universität Berlin

Uli Reich
Freie Universität Berlin

We initially sketch a phonological theory in which the culminativity of word accents acts as only one out of four main functional goals for the configuration of prosodic devices and claim that languages exhibit many differences therein. Thus, in the study of many languages whose prosody is not extensively studied, as is the case for most languages spoken in situations of plurilingualism together with Romance Languages, we need reliable methodologies to determine their particular organization of time, tone, segmental strength and intensity. Using data from a Central Quechua dialect, we propose such a method consisting of a complex pragmatic and metrical annotation in Praat and its statistical exploration in R. We conclude with a discussion of preliminary results and shortcomings to be resolved.

1 The phonological perspective: Competing motivations of prosodic devices

In the study of the phenomena involved in what has been called "accent", "stress" and/or "prominence" in different and partially incompatible terminologies,[1] we

[1]See Beckman (1986) for an impressively well-informed historical overview that sheds some light on the genesis of the terminological confusion and suggests ways out of it. Note that her

Timo Buchholz & Uli Reich. 2018. The realizational coefficient: Devising a method for empirically determining prominent positions in Conchucos Quechua. In Ingo Feldhausen, Jan Fliessbach & Maria del Mar Vanrell (eds.), *Methods in prosody: A Romance language perspective*, 123–164. Berlin: Language Science Press. DOI:10.5281/zenodo.1441341

can draw a line of progress in typological research that puts the universality of the assumption that every phonological word has a single primary accent into question. Early structuralist theory (Trubeckoj 1939) used "main tone" (germ. *Hauptton*) to illustrate the *culminative* (germ. *gipfelbildend*) function in phonological systems as opposed to the *delimitative* and the *distinctive* functions. An accent was conceived of as being "culminative" in the sense that it is the most prominent position in a syntagmatic sequence of hierarchically organized accents. Later, culminativity was developed as a core concept of Metrical Phonology in order to derive "stress" by the hierarchical build-up of prominence in metrical grids (Liberman & Prince 1977: 262; Hayes 1995: 24–25).

In many languages, sometimes called "stress-accent languages", accents are also "accumulative" in the sense that they attract all out of four possible prosodic parameters, namely intensity, duration, pitch and segmental strength. Thus, an accented syllable is believed to show salient pitch events, to be longer, to not reduce vowels, or even diphthongize them (traditionally, the preference for accented syllables for being bimoraic has been coined in Prokosch's Law), to show more complex onsets and codas and to be louder.

A glance at the phonological configuration of tone languages, however, shows that a conception of word accents that accumulate all phonetic instances of strength does not hold cross-linguistically. In tone languages, tones can be associated with many syllables in one word, as Yip (2002) shows in her seminal work. See her example from Chilungu, a language from the Bantu family, in which one tone is associated with many vowels (Yip 2002: 68):

(1) kú-sóóbólól-à ku-soobolol-a

to sort out H

own typological proposal is privative: non-stress languages are languages that are defective with regard to the set of properties that define stress languages. In the following, we will use the term WORD ACCENT to designate the abstract phonological knowledge of one and only one syllable in a prosodic word that is perceived as stronger than all other syllables in that word. In principle, it does imply neither the phonetic cues that may realize it in a given stretch of speech, nor the phonological domain that projects it. Accents can be specified in the lexicon or projected by metrical algorithms that construct alternating *prominence* in a hierarchically ordered metrical grid. In the latter case, word accents are the topmost prominent position in such a grid. The term "stress", as it is used in the literature, is widely synonymous with our notion of "word accent", but implies also aspects of its realization, as in the typological dichotomy of stress-accent and pitch-accent languages. Since we want to keep these notions strictly apart in order to describe their relation precisely, we will avoid the term "stress" wherever it is possible, in spite of its widespread usage in the literature we rely on.

As Yip (2002), among many other scholars working on tone languages, shows convincingly, this one-to-many relation is not the end of the story; many languages also show the inverted relation, associating many tones with the nucleus of one syllable, a situation that is familiar from boundary tones found in many intonational languages.

Duration can also show up without strict association with a word accent. Thus, in Wolof, a Western Atlantic language spoken in Senegal without distinctive tones at the word level, both vowels and consonants show distinctive duration, both for lexical contrasts (2a, b) and the expression of focus (2c, d).[2] Long syllables (bimoraic, hence heavy) are possible in basically every position (3) and in more than one position in the word (4) (examples from Ka 1989; 1994 and Voisin-Nouguier 2002).

(2) a. fat : faat
clean.up kill

b. gën : gënn
better milk

c. ma dem : maa dem
I go [I]$_{Foc}$ go

d. mu dem : moo dem
he goes [he]$_{Foc}$ goes

(3) a. ˈboole
mix

b. teˈraanga
hospitality

c. ˈdajaloò
gather

(4) a. ˈwoowandòo
call.together

b. ˈfeesalukàay
instrument.used.to.fill

c. jiˈgéénubìir
pregnant.woman

[2]Note that the contrasts in (2c) and (2d) appear as focus morphology in the literature. In other cases of the focus system, the morphological contrasts are expressed by segmental modification and addition. See Voisin-Nouguier (2002) and Rialland & Robert (2001) for the full system.

(5) a. ˈtabax
build
b. ˈndaje
meeting

Ka (1989; 1994) claims that in instances of words with light syllables only and in words with one or more heavy (=long) syllable, the first (3c, 5) or the first heavy syllable (3b, 4c), respectively, receives stress. If the heavy syllable occurs after two light ones (3c), it is perceived as having secondary prominence, while primary prominence falls on the initial syllable. In this language, then, duration appears as being as independent from a culminative word accent as tone is in so-called tone languages. As we shall see, the Central Quechua dialects show a distribution of length that comes close to this phonological constellation. Finnish and Latin are European Languages that show distinctive duration independent from word accent, but Latin shows some restrictions that Finnish does not have. German is a quantity language that comes quite close to a "prototypical stress accent language": it has vocalic duration only in "stressed" syllables (Becker 1996). Thus, the independence of duration from word accent shows varying degrees and it is far from clear where we should set the threshold to tell types apart.

Another feature of word accents is that the nuclei of the syllables that bear it, unlike its neighbors, are never reduced, rather often diphthongized and that they show more segmental contrasts than the nuclei of other syllables. Interestingly, exactly this feature has been shown repeatedly as being dependent on what has been called the rhythm type of a given language (Dauer 1983; Auer 1993; 2001; Dufter 2003). In Romance languages, e.g., it holds for European Portuguese, which shows exactly the vocalic reductions and consequently strong restrictions on segmental inventory claimed as general properties of syllables less prominent than the one with the word accent in word rhythmic languages (formerly "stress-timing"). It does not hold as clearly for most varieties of Spanish, Standard Italian or even Brazilian Portuguese, that are taken as instances of syllable rhythmic languages (formerly "syllable-timing", Abaurre & Galves 1998; Reich 2002), but in Brazilian Portuguese, vocalic reduction still occurs more than in Spanish and Italian.

In Spanish, the only property of lexical phonology related to the preference of word accents to be bimoraic is the distribution of diphthongs (cf. */bene'swela/* vs. */beneso'lano/*), while the nuclei of all syllables are fully pronounced in most cases. However, there can be no doubt at all that Spanish does have a word accent that invariably is the locus of major tonal events if they are realized. It is simply less dominant than its European Portuguese counterpart is. Again, we find different

constellations of features that are held to define types. Where should we draw the line?

The last feature we want to mention in this complicated introduction is intensity. Intensity does not seem to play any phonological role at all but as a feature of word accents. Beckman (1986: 160), however, it shows that this property cannot be generalized across languages, since in Japanese intensity is independent from the position of the accented syllable. Intensity, then, also fails to form a universal feature of the strongest syllable in the word.

Given these facts, culminativity rather seems to describe an optional than a universal concept of accents. We hypothesize that culminativity is a functional principle that is counterbalanced by others. Distinctivity is directly antagonistic, as it dissociates time, tone, segmental quality and intensity to enhance the possibilities of paradigmatic contrast for phonological word forms. Thus, the culminativity of word accents grows at the expense of the distinctive potential of prosodic devices. Vice versa, the use of prosodic devices for distinctive functions levels the dominance of word accents, since time, tone, sonority and intensity can be distributed over different positions in the word or phrase. Delimitation may conspire with culminativity towards the overall target of identifying words or phrases in a given chain of speech, but in all systems with non-peripheral word accent, boundary tones, final lengthening and segmental processes like consonantal strengthening and epenthesis of glottal stops are also likely to diminish the phonetic saliency of the syllable bearing word accent. Another core function of accents, absent in structuralist phonology, is rhythmicity, as recognized since the early days of Metrical Phonology (Liberman & Prince 1977; Hayes 1995), but put into a theoretical framework that takes culminativity as universal and thus misses the functional particularities between different aspects of prosodic form. In many languages, the assignment of primary accents does not depend on foot construction (van der Hulst 1999: 72). Rather, the assignment of alternating strengths of acoustic events in time appears to be an independent functional domain of prosody with which lexical accents may, but need not, coincide. The functional target of rhythmicity surely is neither distinctivity, nor culminativity, nor delimitation. To the contrary, it enhances the isochronous distribution of alternating prominence at the expense of all the three functions recognized by the Prague school.

In our view, particular phonologies are organized as instances of decisions between (at least) these major functional goals. Ideal types can be set as abstract possibilities that no language ever reaches because of the competing drives towards the other ends of this space of prosodic possibilities. In the ideal type of

a culminative prosodic system, tonal events, duration, segmental strength and intensity would occur but in the one and only prominent syllable of every word. In the ideal type of a distinctive prosodic system, tones, duration, segmental contrasts and intensity are scattered all over and boundaries between words or phrases are blurred. In the ideal type of a delimitative prosodic system, tones, duration and intensity occur at the boundaries of wordless phrasal chains without prominence. Finally, in the ideal type of a rhythmic prosodic system, we would find isochronally recurring contours of prominence, just as in music (Reich & Rohrmeier 2014). Natural phonologies balance the competing drives towards these ideal types as they fulfil their communicative goals in the variational space of human languages, constrained by universal cognitive principles and the historical traditions of social networks.

In the end, the view we are defending aims at the abolition of dichotomic typologies and pleads for their passage to particular phonological configurations within a polydimensional space defined by competing functional principles. This is very much in line with, while a step more radical than, views on prosodic typologies defended by Hyman (2009; 2014), in his impressive command of facts from the phonologies of many languages in the world.[3] He comes to a conclusion that is not only wise for theory building, but also mandatory for the great empirical endeavor of the study of language: it is better to look at what languages do than to brood upon fictitious universals.

Research into the prosodic phonologies of languages that are not very well known, as is the case for most contact languages of the Romance languages outside Europe, must bear in mind that any constellation of the main prosodic devices may be the case in the language under study. Thus, they must be controlled for independently, but in relation to each other. And we can take nothing for granted.

2 The project and a short overview over the method of data elicitation

The methodological considerations we will expose in the following pages are part of our research project *Zweisprachige Prosodie: Metrik, Rhythmus und Intonation zwischen Spanisch und Quechua* ('Bilingual prosody: metrics, rhythm and

[3] The main difference is probably that Hyman still recognizes the typological validity of concepts such as "stress accent language" and "tone language", while the relevant configurations are only possibilities among others in our prosodic universe.

intonation between Spanish and Quechua'), funded by the Deutsche Forschungsgemeinschaft. The overall goal of the project is the development of a prosodic theory of bilingualism on empirical grounds.

Our data was gathered through fieldwork in Huari, Conchucos, Ancash, Peru during the months of September-October 2015. The elicitation methodology aims at producing semi-spontaneous data, that is, speech produced naturally in dialogues under pragmatic and lexical constraints, in turn influencing prosody. Informants were asked (always in pairs) to play the following dialogical games in which any interaction could only be done orally, gestures were not permitted:

1. Picture-naming. Participants had to name objects shown to them on picture cards.

2. A version of memory. Participants took turns guessing where a certain picture on a card was. The cards had been shown to them for a short time and then flipped over.

3. Map-task (Anderson et al. 1991). Participants were provided with two maps, one with a path drawn between the objects shown on it, the other without the path. They were not allowed to see each other's maps. The participant with the map had to explain the path to the other one, who had to follow it by drawing it on their own map. The maps differed in some of the objects shown. Resulting communicative conflicts had to be resolved orally.

4. Story re-telling. Participant A would listen to a recording of a story (invented by the investigators and spoken for the recording by their local collaborator and advisor, Quechua teacher and native speaker Gabriel Barreta (GB)[4]). They would then tell the story to the other participant B, who had been waiting outside while the recording was playing. After being told the story, participant B would tell the story to one of the interviewers, with the possibility of correction by participant A.

5. A version of "Who am I?". Participant A would be told the name of a person known to both participants, and participant B had to guess the person's identity.

[4]Our deep gratitude and friendship goes to Gabriel Barreta and his family in Huari as well as to Leonel Menacho López and his wife Ana in Huaraz, who have provided invaluable help with local logistic and linguistic questions and without whom our fieldwork could not have been successful.

6. Guessing the contents of boxes. Both participants were provided with closed cardboard boxes and would take turns guessing their contents from just moving the boxes, shaking or weighing them by hand.

The content items of games 1–4 were restricted by means of the props provided, i.e. the cards, the maps and the recordings, to consist mostly of a set of lexical items varying through the possibilities of Quechua syllabic and moraic structure. Hence, the pictures on the cards displayed objects aimed at eliciting words ranging from two light (L) syllables to two heavy (H) and one light syllable, e.g. *tsu.ku* (L-L) 'hut', or *qi.llay.yuq* (L-H-H) 'rich man'. The subsequent games utilized the same lexical items (elicited by means of the same pictures, also on the maps, or the recording) wherever possible. Adjustments to these items were made after a first trial session with our principal local collaborator GB, who gave us local words for several of the metrical constellations that were to be elicited.

Care was taken to have the experiments take place in rooms that were as quiet as local conditions allowed. Participants would play the games in both Quechua and Spanish, going through all the games first in one language, and then the other. Audio recordings of the games were made using a Marantz PMD 670 audio recorder in connection with a Røde NT-1A condenser microphone in 44.1 KHz PCM. In total, excluding the trial with GB, 40 participants (22 females, 18 males, all bilinguals, mean age=22 years) were recorded in 20 sessions, yielding about 7 hours of Spanish and 6 hours of Quechua experimental data. All informants participated voluntarily, gave us their written consent to be recorded and for the resulting data to be published maintaining their anonymity, and were remunerated for their participation.

3 Methodology

3.1 The challenge: determining prosodic constellations on empirical grounds

As far as we know, the determination of positions of prominence in a language where they are unknown has not been studied extensively. In grammatical treatments of understudied (i.e., almost all non-European) languages, accent placement usually is dealt with cursorily: the author describes the positioning of strongest prominence, sometimes differentiating between several acoustic realizations, in a seemingly intuitive manner. Matters such as accent domain, acoustic correlates, function of prominence and interaction with other prosodic phenomena are rarely dealt with in systematically ordered empirical procedures, but attributed

by intuition. Effectively, this means that the describing linguist "hears" accent positions in the language of their informants and generalizes from this audio perception. While we recognize that this was the only methodology available in many cases, we submit that it is not a methodology appropriate for scientific research since it allows for a number of non-trivial descriptive distortions due to perceptual biases on behalf of the describing linguist.[5] If the linguist's first language is one where the lexical accent is culminative and accumulative, as for example in English or German, then we cannot discard the possibility that this phonological background will influence them to a certain degree in their perception of the language they are studying. This will be even more the case if the linguist in question is not a specialized phonologist who might be aware of their own biases in this regard. With the concept of a universal culminative word accent, one is already excluding a large subset of the possible shapes accent systems assume (see, e.g., Hyman 2014 for an overview, and Kügler & Genzel 2012 for a particularly diverging case).

Moreover, in the case of the so-called central dialects of Quechua, we are faced with distinctive length, whose interference with accent placement and realization has never been discussed, and a considerable disagreement even in the existing (impressionistic) literature on accent position, domain and its acoustic correlates. Some of these dialects are described as having primary accent on the penult (Trager 1945 for Huaraz Quechua, Parker 1976 for Ancash-Huaylas Quechua, Adelaar 1977 for Tarma Quechua), others (and sometimes the same by another author) as having it on the initial syllable of the word (Parker 1976 for Huaraz Quechua, Hintz 2000 for Corongo Quechua, Hintz 2006 for South Conchucos Quechua). In most cases, a secondary accent is said to exist on the "other" position, i.e. penult or initial syllable (all of them agreeing at least that no other position is a strong contender), and that their prominence ranking can be reversed under certain morphological, pragmatic or conversational conditions (none of which are agreed upon by any two authors); some recognize a kind

[5]To clarify: our goal is by no means to insult or belittle the efforts of linguists that have done extensive research on otherwise little studied languages. We applaud their endeavors and think that both we and the linguistic community in general are highly indebted to them. However, many of these studies took place several decades ago, when prosodic theory was even more in its infancy than it is today and when, even more crucially, large-scale audio recordings of small languages that could be analyzed appropriately and shared with the academic community were not feasible, due to technical and logistic problems. Many of these researchers had to make do with what they had, and it is no doubt better that they gave an impressionistic description of accent systems in their languages than none at all. Nonetheless we think it is time that with the technical means at our disposal, the methodology to describe prominence systems in these languages should be reevaluated.

of division of labor between acoustic correlates in the realization of prominent positions. Our own visual and acoustic inspection of our own data does nothing to let us decide tentatively in favor of any of the hypotheses suggested in the literature. In fact, it complicates matters, since we often encounter utterances almost entirely devoid of any phrase-internal intonational movement that could reasonably be correlated with lexical pitch accent positions. Intonational movement, if it occurs, seems to respond to a domain above the word and to be severely restricted in its inventory: Utterance-initial rises and utterance- or possibly phrase-final falls and rises are almost exclusively observed. From this inspection, the hypothesis that our Quechua variety does not assign an accent position at the lexical level seems to be at least as probable as any of the suggestions cited above. Another problem is that speakers generally do not receive any education on Quechua and a very traditional one on Spanish, and a bias that accords Spanish greater prestige in academic matters definitely persists. Hence, speaker perceptions on where an "accent" might lie in a given Quechua word and what it might consist of are, if they exist at all, heavily influenced by these social conditions. We therefore thought it necessary to devise a methodology that would help us determine positions of realizational strength from the speech signal, in order to determine without such biases (inherent also in our own perceptions) the nature and domain of regular prominences in the Conchucos variety of Quechua our data is from. In the present study, semi-spontaneous data is used. Hence, phenomena of speech style and individual style are also included in the data. Further research could however easily apply the same method to more controlled data. In the remainder of the text, whenever we say something about "Quechua" without further variational qualification, we mean it to be about the variety of Quechua spoken in Huari, Conchucos, studied by us.

3.2 Goals of the present contribution

While the goal of our overall research project is the development of a prosodic theory of bilingualism on empirical grounds, the present contribution does not yet aim so high. Faced with the conflicting descriptions in the literature and our own data regarding accent placement described above, we think it is important to devise a methodology which arrives at a less biased description of the acoustical data we base our phonological theories on. Given that the question of the domain of accent placement is absolutely vital for any hypotheses regarding the behavior of prosodic domains in a plurilingual context, we take this to be an inevitable first step. Hence, the goal of this contribution is not to say anything about Quechua-Spanish bilingualism, but to provide evidence for the feasibility and usefulness

of its methodology in helping us derive hypotheses about accent placement in a language where this is not known; more specifically, to investigate the possibility by means of acoustic measurements that our variety of Quechua does not assign accent at the lexical level.[6] While all analytical tools and statistic procedures are not new in particular, their complex application to metrical and pragmatic variables has not been endeavored in any empirical prosodic project we know of.

3.3 Methodology – annotation

All the usable Quechua data elicited through the means of the communicative games outlined in section 2 were transcribed and translated by bilingual students of the Universidad Nacional Santiago Antunez de Mayolo (UNASAM) in Huaraz, Ancash, Peru, and morphologically glossed by students of the Pontificia Universidad Católica del Perú (PUCP) in Lima, Peru using ELAN (Wittenburg et al. 2006). All further annotations were done by one of the authors as diligently as possible. In a follow-up investigation on a larger corpus we will also have a part of the corpus annotated by two independent annotators in order to arrive at a measure of inter-annotator agreement.

[6] There is a substantial body of literature on Spanish-Quechua bilingualism with regard to all levels of linguistic description, including prosody, such as O'Rourke (2005; 2007; 2008; 2009); Muntendam (2010); O'Rourke (2010; 2012); Muntendam (2012a; 2012b); van Rijswijk & Muntendam (2014); Muntendam & Torreira (2016); Muysken & Muntendam (2016). All of it is concerned with Southern or Ecuadorian varieties of Quechua (that are more closely related to each other than to Central Quechuan varieties), where accent placement seems not to be problematic. Due to the fact that in other systematic domains of the language, such as morphology and segmental phonology, the Southern and Ecuadorian varieties are different enough from the Central ones as a whole (not considering their considerable internal variation) to be judged mutually unintelligible, we are hesitant to just assume that the findings in the literature on the interaction between Southern Quechua and Spanish intonation, e.g., can easily be applied to our variety. Mountainous regions notoriously harbor enormous variation in a relatively small geographic area, that this applies also to the realm of prosody is by now well-known at least from the case of the varieties of Basque, where more or less every prosodic parameter regarding accent placement can be found (cf. e.g. Hualde et al. 2008; van der Hulst 2010; Aurrekoetxea et al. 2012). That the varieties of Quechua vary with regards to their accent placement to a considerable degree has also been recognized, see Wetzels & Meira (2010). Our goal here is only to provide a better basis for a description of the accent placement in the variety of Quechua we are studying. Only in subsequent studies will we hopefully have to say something about bilingualism.

3.3.1 Syllables

Using a corrected version of the transcription as basis, we built on it by adding a syllabic annotation on a map-task by two male speakers (FB03 and WB04) in Praat (Boersma & Weenink 2017). An example of what the syllable annotation looks like can be found in Figure 1. Tiers 4 and 5 in the Praat textgrid are reserved for syllable annotation (one tier for each speaker). Annotation boundaries were aligned as closely as possible with corresponding beginnings and endings of segmental material. A transcription system was used that aims at grouping together segments belonging to the same relevant class in Quechua.

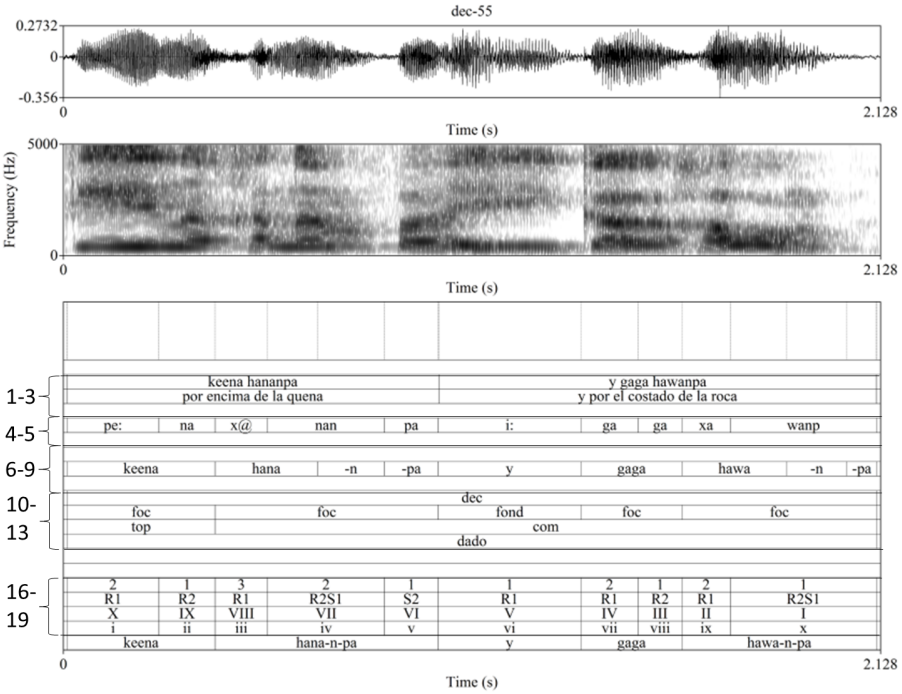

Figure 1: Waveform, spectrogram and text grid of the utterance *keena hananpa y gaga hawanpa*.

This was done to avoid illusions of perfect phonetic transcribability of spontaneous data on the one hand and to facilitate categorizations using these classes in the analysis on the other. Thus, in the example in Figure 1, what is morphophonologically written in (6)

(6) ke:na hana-n-pa i gaga hawa-n-pa
flute above-3S.POSS-GEN and rock below-3S.POSS-GEN
above the flute and below the rock

is transcribed as in (7).

(7) <pe:|na|x@|nan|pa|i:|ga|ga|xa|wanp>

Voiceless plosives are transcribed with <p>, back fricatives [x ɣ h] with <x>, all nasals with <n>, reduced vowels with <@>, etc. Note that a phonetically somewhat more accurate rendering of the part transcribed as <x@nanpa> would be [xənampa]. Another remark needs to be made regarding syllable structure. Note that in the transcription of /hawa-n-pa/, as opposed to /hana-n-pa/, <wanp> has been grouped as one syllable, with a syllable structure of CVCC, something which has not been described in the existing phonologies of Quechua varieties (cf. Parker 1976 where a maximal syllable is either CVC or CV:). This is because of the complete elision of the vowel /a/ of the genitive marker *-pa* in this case, as can also be seen on the spectrogram in Figure 1, where there is no visible release of the plosive in the part corresponding to <wanp> as opposed to the plosive in the part corresponding to <nanpa>. This elision of vowels (as well as their sometime reduction) is a very frequent phenomenon in our variety of Quechua, occurs both utterance-finally and –medially and is not restricted to this particular morpheme (which can also be realized fully). Note that we transcribed what was spoken and not any assumed underlying forms.

Tiers 6–9 in the textgrid are used for a morphological annotation and glosses (two tiers for each speaker). Whereas the syllabic transcription is a reduced transcription of the actual realization, the morphological annotation represents morphemes in a form close to standard descriptions. Interval boundaries in the morphological tiers were made to coincide as closely as possible with the corresponding sound changes in the speech signal, however this was not always possible: a frequent example involves the copula *ka-* together with the progressive *-yka-*, as in

(8) ka-yka-n
COP-PROG-3S
'S/he is in the process of being/having.'

Here the usual (but not the only) realization is [ke:kan], arising from a regular process of monophtongization in many central Quechua varieties, thus making it impossible to determine where in the speech signal the boundary between the

root and the progressive suffix lies. In such cases, labels on the morphological tier received a <?> after or before the connecting hyphen in the annotation.

3.3.2 Information structural annotation

As can be seen from Figure 1, the next four tiers in the annotation textgrid are used for information structural pragmatic categories. Tier 10 annotates the speech act, with the categories DEC for declaratives, PREGQ for queries, PREGCH for checks and IMP for imperatives. Tier 11 is used for the annotation of focus-background structure, with FOC for focus and FOND for background. Tier 12 annotates topic-comment structure, with TOP, topic, and COM, comment; tier 13 annotates givenness in the discourse with the categories of *dado* 'given', and *nuevo* 'new'. These annotations were made based on judgments about what role the parts of the utterance in question played in the discourse, not based on the presence or absence of morphological markers that have been described as encoding information structural meaning in Quechua (cf. e.g. Wölck 1972; Weber 1986; Muysken 1995; Gómez Rendón 2006). Building on standard approaches to pragmatics and information structure (Austin 1962; Searle 1969; Chafe 1976; Bolinger 1989; Rooth 1992; Grice & Savino 1997; Baumann 2006; Krifka 2007, among many others) we elaborated the following set of key notions:

- A DECLARATIVE is adding a proposition to the common ground, regardless of whether this proposition is claimed by the speaker to be true (asserted) or not.
- A QUERY asks for information needed to complete a proposition, whether regarding its components (constituent question) or its truth value (polar question).
- A CHECK asks for confirmation that a proposition in the common ground should be considered to be true in the relevant world of discourse.
- An IMPERATIVE represents a command by the speaker (to the hearer) that a certain state of affairs in the world should be changed so as to conform to a proposition.
- In FOCUS are those parts of an utterance to which alternatives are saliently evoked and discarded; those which make a difference to propositions about states of affairs already in the common ground.

- Correspondingly, in the BACKGROUND are those parts of an utterance to which alternatives are not evoked, that are not in focus; background is complementary to focus.
- TOPIC is that part of a discourse that is currently being talked about or advanced by one of the participants in the discourse to be talked about henceforth, it is the frame of reference under which comments are to be understood.
- COMMENTARY is that which is talked about the topic; it is information that is added to the common ground regarding the topic.
- GIVEN are referents that are active in the discourse, whether they have been activated by linguistic or extralinguistic means.
- NEW are referents newly introduced into the discourse, they become active through their first (linguistic or extralinguistic) mention.

These definitions are not without their problems: in particular, there exist certainly many more types of speech act than just the four defined above; new and given are categorical impositions on a multitude of states of discourse activation that probably should be thought of as forming a graded scale; different types of focus such as information focus and contrastive focus can and have been argued for (for a proposed hierarchy of them see Féry 2013), and it is probably useful to further divide background into tail and link, as proposed by Vallduví (1992). However, this part of the annotation is complex enough as it is, and we therefore decided to restrict ourselves to the above-mentioned categories, believing that they should suffice for the time being for the purposes of determining the relevance of information structural categories for the realization of prosody in our data.

These categories encode related but clearly distinct notions. Interactions may arise, e.g. when a new topic is introduced into the discourse it will often be in focus because it is then that it is contrasted with other alternatives. However, once it is introduced and well known, it is frequently not mentioned anymore, but the comments made about it are still divided between focus and background. This is only to make the point that the two are neither complementary nor the same, nor are any of the other categories defined above and annotated in separate tiers the same. We hold that it is not possible to further reduce these categories, such as they are defined above, without severely limiting one's descriptive power and hence the scope of phenomena one would like to explain.

It has only been possible to annotate these information structural categories in such comparative detail and on the basis of only discourse pragmatic considerations because the annotators both also designed the experiments that produced the semi-spontaneous data (controlling a large percentage of the content words metrically) and were present during the course of the experiments as silent participants. Thus both the states of affairs that are talked about as well as the (linguistic and extralinguistic) discursive progression of events are well known to us, and our annotations (of course remaining interpretations to a certain degree, but this is valid for all annotations) are anchored in and informed by these facts.

3.3.3 Positional annotation

Tiers 16–19 in Figure 1 demonstrate what we call positional annotation. In order to create the intervals, a Praat script first cut the entire map task recording with its textgrid into chunks at the speech act grid that correspond roughly to conversational moves and are surrounded by (small) silences. Another script then detected in which of the two syllable tiers, annotating the two speakers respectively, there were more intervals and made four empty copies of its intervals in tiers 16–19, thus selecting only the more dominant speaker for each such chunk to be analyzed. In most cases, this yielded sound-textgrid pairs without any speaker overlap. In those cases where there was speaker overlap, the parts where the less dominant speaker was speaking were excluded from the analysis. The intervals in tiers 16–19 were then labelled as follows, creating the positional annotation:

Tier 16 annotates syllable position in the word counting from its right edge, using Arabic numerals (1, 2, 3, 4, etc.). Thus, in the example given and starting from the right, the syllable <wanp> receives in its corresponding interval on tier 16 a <1>, <ha> a <2>; the right <ga> from *gaga*, because it is a separate word, again receives a <1>, and to its left the other <ga> a <2>. "Word" here refers to the morphological word in Quechua whose structure all grammars agree upon, i.e. consisting of a root plus several suffixes. Theoretically, due to the agglutinative nature of Quechua, any number of suffixes could be attached to a root; in practice, the furthest a syllable was annotated as being removed from the right word boundary in this corpus was 6, and 3 was not often exceeded. The implicit assumption here is that the domain of the morphological word is largely isomorphic with a relevant prosodic domain, e.g. the prosodic word in Quechua, if it exists, although this is in fact unknown. Note that what is annotated here are syllables as defined in the part on syllable transcription and transcribed in tiers 4 and 5, not morphemes.

Tier 17 annotates syllable position in the word counting from its left edge as well as morphological category (root or suffix). The labeling consists of either "R", for root, or "S", for suffix, plus an Arabic numeral indicating position and whose counting is reset at the border between root and suffix. To clarify using the example from figure 1: Starting from the left, <pe:> and <na> are labelled <R1> and <R2> respectively, because they are both part of the root of the verb. Proceeding, <x@>, as the first syllable of the next word, gets labelled <R1> again, but <nan>, because it consists both of the second part of the root *hana-* 'above' and the first suffix *-n* "3rd singular possessive", is labelled <R2S1>. With this twofold annotation, it is possible to examine both position from left-edge word boundary and whether a syllable is part of the root or the suffixes of a word as influencing factors on prosody in the later analysis.

Tier 18 annotates syllable position in the whole utterance counting from its right edge using large Roman numerals (I, II, III, IV, etc.). The utterance is here defined as consisting of what is labelled as one conversational move between (short) pauses, hence everything one speaker says "in one go". This may correspond in many cases to a prosodic phrase, if only to postulate a prosodic domain distinct from the word in Quechua. Whether this phrase is in itself composed of further smaller phrases that are not isomorphic with the (postulated) prosodic word in Quechua is not known (see Grice et al. 2000 for discussion of the concept). To look at the example, the interval on tier 18 corresponding to the rightmost syllable <wanp> is labelled with a <I>, from there the numbering increases rightward until the interval corresponding to the leftmost syllable, <pe:>, which gets labelled <X> for being the tenth syllable in the entire phrase counting from the right.

Tier 19 works exactly as tier 18, only counting from the left edge of the utterance/phrase and with the numbering being done in small Roman numerals (i, ii, iii, iv, etc.). Thus, the leftmost syllable <pe:> here gets labelled <i>, and <wanp> at the right edge <x>, for being the tenth syllable in the phrase if counting from the left.

Proceeding in this way has several advantages: Only by using four tiers for positional measurements can we exactly observe and quantify prosodic behavior at both the left and the right edge of two domains, that of the word and that of the phrase. Note that none of the tiers annotates information already given in another tier; since both word length and phrase length are highly variable, the left-counting tiers can only give precise positional information at the left edge of the domain, and vice versa for the right-counting ones. This procedure incorporates standard assumptions in every theory of metrical phonology (e.g. Liberman

& Prince 1977; Hayes 1995; van der Hulst 1999).[7] Since it is so far unknown which prosodic domains play a role in prominence assignment in Quechua, it would be inadvisable to neglect one of these domains by not annotating for syllabic position in it. On the other hand, if the later analysis shows consistent results e.g. in the prosodic behavior of syllables counted from the right edge of the phrase but not the word, it will be possible to conclude that this domain definitely does play a role. We think that the domains annotated as they are here are the best possible candidates so far for playing a role in influencing prosodic phenomena and only after the final analysis will it be possible to see where they need to be further refined.

To complete the description of the annotation process, there is also a derivative tier of the morphological tier, tier 20, that annotates word-length elements by copying the tier of the morphological annotation and leaving only those boundaries that are to the left of the beginning of the root of a word (using information from tier 17). A script in Praat was written for that and the results checked. This helps to take important measurements in relation to word length, an information that wasn't contained in the textgrid up to that point, as will be seen in the next section.

3.4 Measurements

A Praat script was written that used the information encoded in the annotation textgrids detailed above. Per annotated syllable, it was used to extract the relevant annotational information from the tiers in the textgrid described above, i.e. what word the syllable belongs to, which one of the two speakers is uttering it, what information structural categories it is annotated for, what position according to the positional annotations it has, etc. Acoustic measurements per syllable were also taken by it from the sound files in the corpus. The measurements taken can be divided into two kinds: absolute and relative.

3.4.1 Absolute measurements

We extracted measurements of F0, intensity and duration, all of which have been shown to variously play a role in the encoding of prominence. In particular, the absolute measurements taken were:

[7] If our method were to show that no kind of prominence is computable from the edges we would have to assume prominence as a diacritic in the phonological word in Quechua. This is hard to believe. Rather, we expect some aspects to be derived metrically and others to be lexically fixed.

1. Syllable duration in ms
2. Mean F0 per syllable in Hz
3. Minimum F0 per syllable in Hz
4. Maximum F0 per syllable in Hz
5. Position of minimum and maximum F0 within the syllable
6. Pitch range within the syllable in Hz
7. Mean intensity per syllable

From the measurements for F0 minimum and maximum, the script further created a categorical measurement of F0 movement for each syllable: if minimum and maximum were at least 30 Hz and 100 ms (about one standard deviation of the syllable duration in the sample) apart from each other, the syllable would be assigned one of the labels "rising" or "falling", depending on whether the movement was from minimum to maximum or the other way round; if these criteria were not met, the syllable was assigned the label "level".

3.4.2 Relative measurements

The relative measurements are based on the deliberation that prominence by definition is a relative concept. It is impossible to make a statement about the prominence of a syllable just from knowing that e.g. it has a certain mean F0 value or even that it has a large pitch range, without a comparison with the corresponding values of other units, i.e. that of adjacent syllables or the mean value of the word the syllable is a part of. That is exactly what the relative measurements do (for a similar approach, see Pamies Bertrán 1996). For most of the absolute measurements, the script also produces relative values that serve as a comparison with the corresponding acoustic parameter on three levels: that of the adjacent syllable (left and right if they exist, i.e. if the syllable isn't itself at a domain edge), that of the word, and that of the phrase (each as defined within the annotation method described above). These are the relative measurements obtained per syllable:

1. Syllable duration divided by average syllable duration within the phrase
2. Syllable duration divided by average syllable duration within the word

3. Syllable duration divided by duration of the left-adjacent syllable (if it exists)
4. Syllable duration divided by duration of the right-adjacent syllable (if it exists)
5. Mean F0 of the syllable divided by mean F0 of the phrase
6. Mean F0 of the syllable divided by mean F0 of the word
7. Mean F0 of the syllable divided by mean F0 of the left-adjacent syllable (if it exists)
8. Mean F0 of the syllable divided by mean F0 of the right-adjacent syllable (if it exists)
9. Pitch range of the syllable divided by pitch range of the left-adjacent syllable (if it exists)
10. Pitch range of the syllable divided by pitch range of the right-adjacent syllable (if it exists)
11. Mean intensity of the syllable divided by mean intensity of the phrase
12. Mean intensity of the syllable divided by mean intensity of the word
13. Mean intensity of the syllable divided by mean intensity of the left-adjacent syllable (if it exists)
14. Mean intensity of the syllable divided by mean intensity of the right-adjacent syllable (if it exists)

A few remarks need to be made regarding these measurements. In general, the script was written so as to recognize when a syllable had no left- or right-adjacent syllable, so it wouldn't take the corresponding relative measurement. Then, because they are obtained by dividing the value of a parameter for the syllable that is being investigated by the corresponding value of another unit, all of these relative measurements are naturally grouped around the value 1 in the sense that if they are larger than 1, it means that the value of the syllable in question is higher than that of the unit it is compared with; if it is below 1 it is lower than that of the other unit in comparison; if it is exactly 1 the two values

compared are the same. These values increase exponentially, therefore their log-10s were taken so that they are now grouped around 0, increase linearly and statistical measurements such as means can be applied to them.

Note that we originally intended to also compare formant measurements between syllables in order to look at vowel reduction as a correlate of non-prominence. However, implementing this would not be straightforward, as only phonologically same vowels that are adjacent could be reasonably compared, and then only via a derived measure comparing e.g. distance between F1 and F2, as an indicator of centralization of the vowel. However, centralization is not the only way of reducing vowels,[8] so this would still be an insufficient procedure. We decided to exclude vowel quality from the scope of this preliminary study and to perform a detailed analysis of it later.

3.5 Putting things into interaction

A total of 1019 syllables of the pilot corpus were annotated and measured in the way described above. 26 of these had to be excluded because of overlap between speakers, reducing the number of syllables that can be used in the analysis to 993. After applying the steps described above, there are now per syllable up to 43 numerical (ratio and interval) variables obtained through the acoustic measurements and 12 categorical (nominal and ordinal) variables obtained by extraction of the relevant annotation information from the textgrid. A further categorical variable, syllable type (C, CV, V:, CVC etc.), was derived from the syllable annotation by using regular expressions in R (R Core Team 2016). Now, the purpose of this approach is to bring to light the effect each of the linguistic categories encoded in the annotation has on realizational strength, which is supposed to encode phonological prominence. For this purpose, the data was imported into R, so that the measurements could be plotted in dependence on the categorical variables, either in isolation or in conjunction. There are two important points to consider. The first is that in our pilot corpus, we do not have a balanced sample with regards to frequency of occurrence of the variants the categorical variables

[8] Note that in two studies on (unstressed) vowel reduction in Cusco Spanish (Delforge 2008) and Cusco Quechua (Delforge 2011), the phenomenon is described as existent and wide-ranging but consisting phonetically of instances of vowel devoicing with no apparent centralization. While this might well be the case, the fact that Cusco Quechua and the Central Quechua varieties are so different in many other respects means that we are unwilling to simply transfer these conclusions to our case. Besides, Delforge (2011) never connects vowel reduction in Cusco Quechua with any prosodic properties of the language, so that the relation between vowel reduction and stress or accent in Quechua remains entirely unexplored so far.

may assume. That is, if we were to compare the mean values of the rightmost syllable in the word with that of the fifth rightmost syllable in the word just like that, the results would be skewed, because there are just 15 syllables in the corpus that are annotated as being in the fifth rightmost position of the word, whereas there are 410 syllables in rightmost position within the word. The same would happen when comparing declaratives with imperatives; while there are more than 600 syllables annotated as belonging to the former category, there are only 15 for the latter. Differences in variance between the two groups to be compared would result more from the differences in sample size than from any inherent properties of the underlying populations. Ideally we would need a much larger sample, so even rarer variants and conditions would still have enough occurrences to allow for a meaningful comparison. Since this is not feasible with the sample size we have, a preliminary solution must be to exclude rare variants and to only work with those that have a reasonable number of occurrences in the dataset. The second point is related, but of a more linguistic nature. When comparing measurements, again, between the rightmost, the second rightmost and the third rightmost syllable of a word, the results will be of little value if care isn't taken to make sure that the second and third rightmost syllables aren't sometimes initial syllables, a fact that might influence prosodic realization. In other words, it makes more sense to compare syllables belonging to words of the same length.

3.6 The realizational coefficient

One approach that is interesting to pursue exploits the nature of the relative measurements taken. Since they are all (log-10s of) ratios of a parameter measured for one syllable to that of an adjacent (left or right) syllable or larger unit, they are all at the same scale and therefore comparable. They can also be used in combination. Adding together relative values for one measurement with those of another and dividing by their number, we get average ratios per syllable of several parameters at once. For example, we can get an overall realizational value for the ratio of one syllable to its adjacent syllables by taking the relative values of syllable duration divided by that of its left- and right-adjacent syllables, as well as those for mean F0, pitch range and intensity, adding them together and dividing them by the number of values added (eight in this case). Since these values express relative realizational strength of one syllable over others, we may name the result (somewhat tongue-in-cheek) the overall realizational coefficient(s). With the help of these coefficients we can determine the relative realizational strength of any syllable position in our data, as well as (by comparison with the singular rel-

ative measurements) its contributing factors. Thus it is possible to show not only why a certain syllable position is strong when viewed e.g. at the word level, and what parameter is responsible for this, but also why it might not be very salient at another level (e.g. the phrase level), for example because the effect one parameter has at one level is overruled by that of another at another level. It needs to be pointed out that the same coefficients cannot always be used equally. Syllables at phrase edges have empty pauses to one of their sides, and the Praat script did not take relative measurements for them towards that side, see §3.4.2. For that reason, when looking at phrase-level measurements, the values comparing syllables with average syllable values in their phrase were used. In theoretical terms, using these realizational coefficients will help us to disentangle the interaction of the different means of creating prominence at different levels.

4 Preliminary results

In the following, some of the findings resulting from looking at the measurements at the different syllable positions under some of the information structural conditions will be presented and discussed. The purpose is here to explore the capabilities and limitations of the method described for creating profiles of realizational prominence in the data. In order to gain a comprehensive overview of the multidimensional dataset, hundreds of combinations of conditions were examined in different ways. We can only report on a few of the ones that seem to be most promising for future exploration. Since the sample is not balanced, we decided not to use any inferential statistics in order not to create false impressions of the general validity of our results unsuitable to the exploratory nature of this study.

4.1 Prominent syllables: Penultima

One important overall finding that derives from several ways of looking at the data is that the penultimate syllable is indeed prominent in the sense that something is happening there, but that it isn't at all certain that the prosodic domain of which this is the penultimate syllable is really isomorphic with the word as defined here. To explore this finding in more detail, let us proceed from looking at overall values to more individual ones:

Figure 2 shows barplots of the median values obtained for the first four syllable positions in the word (we left out the fifth and sixth due to small token size and better visibility) counted from the right word boundary, for the whole dataset

used. The ends of the bars indicate the (medians of the) ratio of the parameter measured for the respective syllable position, divided by its immediate neighbours to the left and right (where applicable). The parameters used are mean F0 per syllable (red), pitch range (orange), duration (blue), intensity (green) and all four combined and divided by their number (black, from here on called "the left-right overall coefficient", see above). Considering the left-right overall coefficient, the penultimate syllable of the word obtains median (0.0081) and mean (0.0072, sd=0.1197) values only slightly above zero (meaning that at least half of its values are better than those of its adjacent syllables). Almost all other syllable positions also have median values slightly above zero, and their quartiles (see Table 1) range very far both below and above zero. Hence the penult is not particularly *more* prominent in this regard than other positions. The values for the left-right overall coefficient are very small here, even compared to the already small scale used in this figure.[9] This reflects the fact that the parameters it is composed of do not show a uniform pattern, i.e. do not unequivocally indicate a prominence for the penult at all. For example, the penult actually has a median slightly below zero for the parameter of duration, and at zero for the parameter of pitch range. Duration seems to instead favour the anteantepenult and pitch range the ultima, and also for mean F0, the median values for the penult do not stand out particularly from those of the antepenult. Hence, the picture we get from this very first approach to the data is less than unified, and does not seem to indicate a state of

[9] It has to be borne in mind how these graphs are to be read: the values on the y-axis are logarithms of base-10 (log-10) of the ratio of a value obtained in a parameter for a syllable in the indicated position divided by the values obtained for the same parameter by either its adjacent syllables or the larger unit (word/phrase), thus representing relative realizational strength of the syllable in that parameter (see §3). A value of zero on the y-axis therefore means that the syllable is equally strong in the parameter as the unit it is compared with, positive values mean that it is realizationally stronger relative to the compared unit, negative values mean that it is weaker. A log-10 value of 0.1 corresponds to a ratio of about 1.259:1 (meaning that the value for the syllable is about 25 % higher than that of the unit it is compared with), a log-10 value of 0.04 to a ratio of 1.096:1 (a little less than 10 % higher), a log-10 value of 0.02 to a ratio of 1.047:1 (a little less than 5 % higher), a log-10 value of 0 to a ratio of 1:1, a log-10 value of -0.02 to a ratio of 0.955:1 (a little less than 5 % below the value for the compared unit), a log-10 value of -0.04 to a ratio of 0.912:1 (a little less than 10 % below), and a log-10 value of -0.1 to a ratio of 0.794:1 (a little more than 20 % below), and so on, to give an example. The barplots indicate the median value for each syllable position. Each individual measurement obtained above zero means that this *individual* syllable is more realizationally strong in the parameter measured than its *specific and individual* neighbors. A median above zero thus means that this is the case for at least half of all syllables in this position. Because of the lower quartiles reaching below zero (see Table 1) and the differing token sizes per syllable position, there is no contradiction with several positions having their medians above zero: this only means that there is a smaller percentage of syllables that are less realizationally strong than their individual neighbors.

affairs in which all acoustic parameters unite in order to make a single syllable in the domain of the prosodic word stand out from all others. Taking a look at specific conditions, word lengths and the individual realizational parameters will hopefully provide more detailed insight.

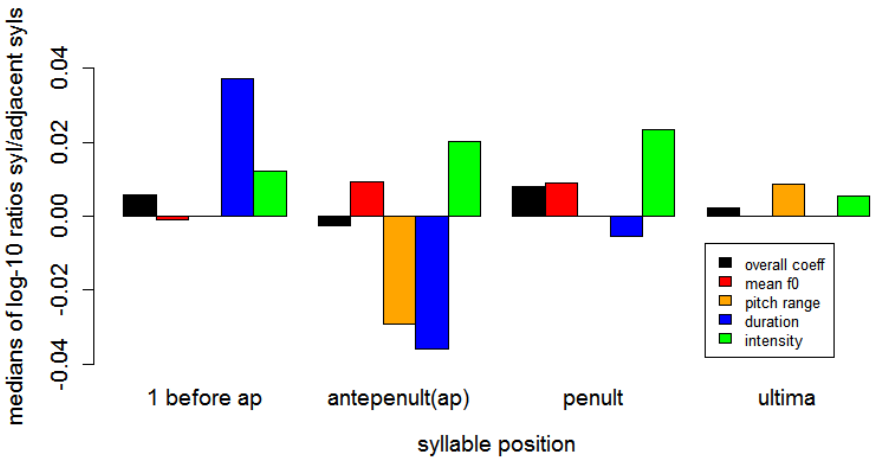

Figure 2: Barplots comparing the medians of 5 log-10 ratios obtained by dividing the value for the syllable by that of its adjacent syllables in the word, ordered according to syllable position in the word from the right word boundary: the overall (black), mean F0 (red), pitch range (orange), duration (blue) and intensity (green) left-right coefficients for the entire sample. N (1 before antepenult) = 49; N (antepenult) = 136; N (penult) = 281; N (ultima) = 255

4.2 Word length

As explained above, by controlling for word length we eliminate possible effects of left-edge phenomena. If only words of the same length are compared, we can observe not only the right, but also the left edge of the word. The syllable position that is leftmost in the graph is now also the leftmost syllable in the word. Additionally, token sizes for each syllable position are now equal (almost, because of elimination of syllables for reasons such as speaker overlap or being at the edge of the phrase). We can thus create more reliable prominence profiles for each word length. Unfortunately, in this sample, this kind of reduction also means that we can only effectively observe words from lengths 2–4 (see token sizes given in the descriptions for the figures).

Table 1: Values for relative parameters in words of all lengths, counting from the right word boundary (see Figure 2)

type of coefficient	syllable position in word	lower quartile	median	mean	upper quartile	sd	N tokens (syllables)
overall	3 before antepenult	−0.09176	−0.06776	−0.06776	−0.04375	0.06790292	2
	2 before antepenult	−0.05519	0.02811	0.01336	0.05518	0.08698023	11
	1 before antepenult	−0.04438	0.00570	0.01235	0.07184	0.10113552	49
	antepenult	−0.061365	−0.002598	−0.007605	0.048850	0.10058366	136
	penult	−0.059410	0.008110	0.007206	0.078412	0.11970251	281
	ultima	−0.061826	0.002242	0.003760	0.083105	0.12473560	255
mean f0	3 before antepenult	−0.032162	−0.021081	−0.021081	−0.010000	0.03134178	2
	2 before antepenult	−0.003155	0.007354	0.007931	0.027405	0.04696297	11
	1 before antepenult	−0.0319673	−0.0009873	−0.0057637	0.0106195	0.04610146	49
	antepenult	−0.009671	0.009344	0.006082	0.026019	0.05388791	136
	penult	−0.007283	0.008938	0.006025	0.031923	0.05610093	281
	ultima	−0.025128	−0.000083	0.002712	0.024209	0.08117007	255
pitch range	3 before antepenult	−0.37810	−0.17418	−0.17418	0.02975	0.5767830	2
	2 before antepenult	0.0000	0.1431	0.1552	0.2426	0.3651206	11
	1 before antepenult	−0.19496	0.00000	0.07306	0.28384	0.4362566	49
	antepenult	−0.35170	−0.02925	−0.06869	0.15508	0.4262779	136
	penult	−0.22749	0.00000	0.02314	0.28935	0.4921704	281
	ultima	−0.231225	0.008477	0.005969	0.316592	0.5207739	255

Table 2: Values for relative parameters in words of all lengths, counting from the right word boundary (see Figure 2)

type of co-efficient	syllable position in word	lower quartile	median	mean	upper quartile	sd	N tokens (syllables)
duration	3 before antepenult	−0.1461	−0.1414	−0.1414	−0.1368	0.01308885	2
	2 before antepenult	−0.19105	−0.05224	−0.07615	0.06760	0.17805208	11
	1 before antepenult	−0.120304	0.037103	0.002385	0.151268	0.20698473	49
	antepenult	−0.16943	−0.03586	−0.01976	0.11743	0.20664683	136
	penult	−0.169099	−0.005398	−0.016198	0.149812	0.24136864	281
	ultima	−0.158923	0.000000	−0.003881	0.154910	0.23593383	255
intensity	3 before antepenult	0.010245	0.012472	0.012472	0.014700	0.006300656	2
	2 before antepenult	−0.007806	0.016768	0.006260	0.022976	0.025963067	11
	1 before antepenult	−0.004430	0.012249	0.009791	0.023763	0.023103567	49
	antepenult	0.004583	0.020092	0.018961	0.035845	0.027377440	136
	penult	0.009766	0.023370	0.022498	0.038142	0.025619751	281
	ultima	−0.010866	0.005356	0.001503	0.018054	0.027426937	255

Both two-syllable and three-syllable words do not especially indicate prominence for the penult (see Figure 3, comparing the left-right overall coefficient for words of length 2 (black), 3 (red) and 4 (blue)).

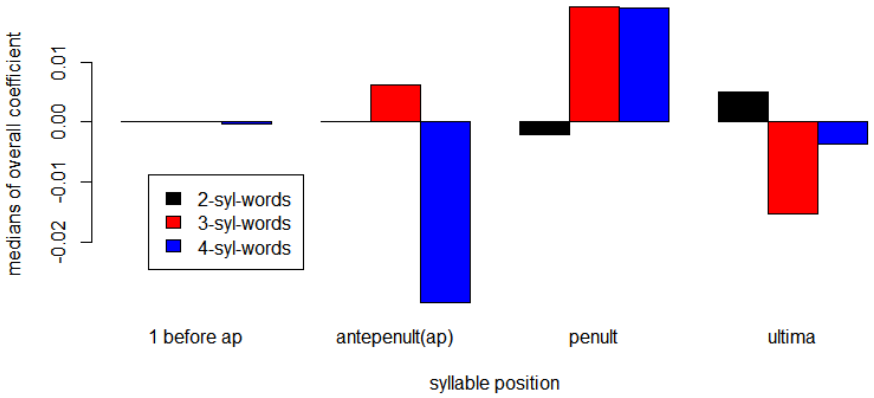

Figure 3: Barplots comparing the medians of the overall left-right coefficient for words of syllable length 2 (black), 3 (red) and 4 (blue), ordered according to syllable position in the word from the right word boundary. For 4-syllable words: N (1 before antepenult) = 31; N (antepenult) = 43 (4-syl words); N (penult) = 44; N (ultima) = 28. For 3-syllable words: N (antepenult) = 71; N (penult) = 110; N (ultima) = 72. For 2-syllable words: N (penult) = 105; N (ultima) = 116

The medians for all positions are more or less the same in 2-syllable words, and in 3-syllable words the median of the penult is the highest, but not by much. Not much indication, either, for a stronger realization of the initial syllable. However, in 4-syllable words, the picture changes (see Figure 3). Here, the median of the penult is visibly higher than that of its surrounding positions.

The initial syllable is also stronger in its realization than the ultimate and antepenult. This would give some support to the proposal of a primary prominence on the penultimate, and a secondary prominence on the initial or every second syllable from it (which of the two cannot be determined here), or one where prominence is assigned metrically to every second syllable in a unit. It has to be kept in mind however, that the ratios obtained here are again very small overall, indicating differences in the realizational strength of the syllables of at most 10–15 %.

4.3 Different parameters

We will now explore the factors contributing to the values of the overall realizational coefficient, i.e. the three acoustic parameters it consists of, duration, F0 and intensity. We will see how they support the finding of the penultimate syllable receiving word-level prominence.

4.3.1 F0

There are several ways in which F0 can reasonably influence realizational strength. Leaving other things aside, an interesting difference arises between the relative measurements of mean F0 per syllable and pitch range per syllable divided by their adjacent syllables. See Figure 4 showing the values for pitch range (dark blue) and mean F0 (red) per syllable, both in words of length 4.

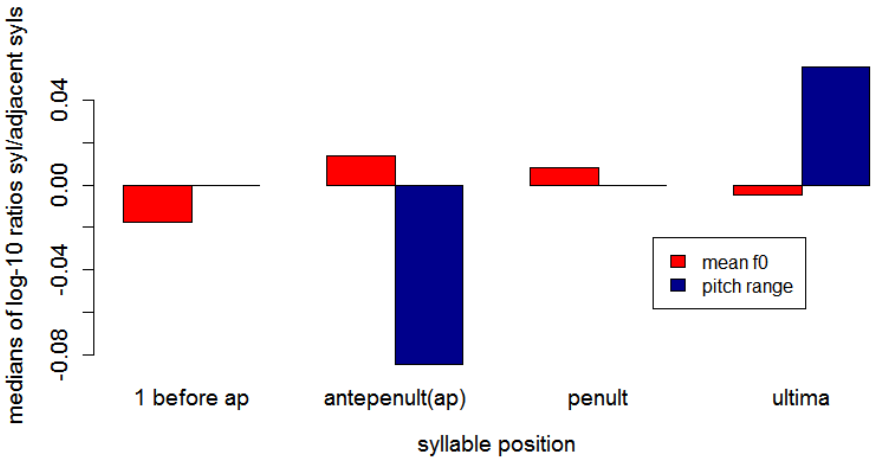

Figure 4: Barplots comparing the medians of the left-right pitch range (dark blue) and mean F0 (red) coefficients for words of syllable length 4, ordered according to syllable position in the word from the right word boundary. N (1 before antepenult) = 31; N(antepenult) = 43; N (penult) = 44; N (ultima) = 28

While the values for pitch range seem to display an alternating pattern, with the ultima realizing the largest range in comparison with adjacent syllables, the antepenult the least, and the penult and initial syllable being more or less equal, the values for plain F0 form a sort of arc, with low values at the edges and high ones in the two middle positions. These two results are what would be expected

if the actual F0 pattern was that of a rise on the initial syllable, a plateau on the intervening one(s), and a drop on the penult (and/or ultima, see below).

This ties in with observations we made inspecting the corpus individually: the main tonal movements seem to be a fall on the penult, often combined with a severe reduction of the last syllable, and a less pronounced initial rise. At the phrase level, a similar pattern manifests, of a (phrase-)initial rise, slow downtrend throughout the phrase and additional movement on the penultimate and last syllable (see Figure 5 as a good example of the overall persisting pattern) – note that there is considerable variance on the values of the last two syllables of the phrase, likely due to additional phrase-final movement used to encode assertions and questions, or finalizations and continuations. Comparing the F0 values for the phrase and word levels, there is an indication that a considerable falling movement often takes place over the last three syllables - but this is the case for both words and phrases, so it is hard to tell whether it is indeed a word- or a phrase-level effect.

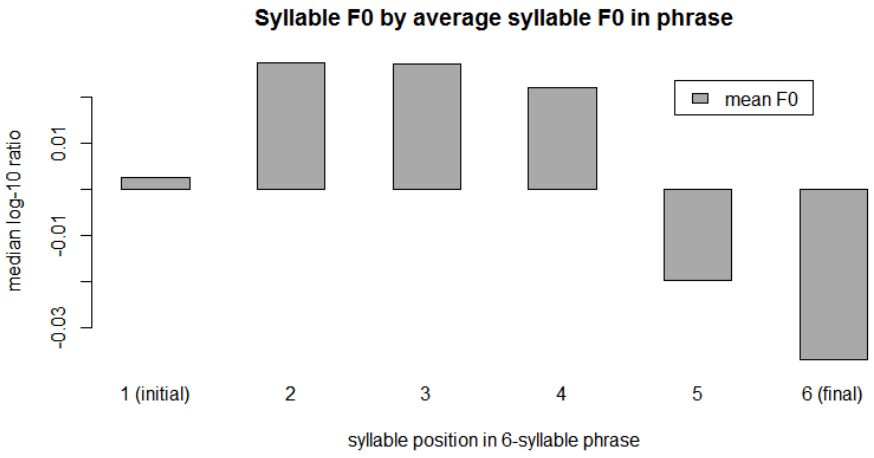

Figure 5: Barplots of median log-10 ratios of mean F0 per syllable divided by average mean F0 per syllable in phrases of syllable length 6, ordered according to syllable position in the phrase from left to right. N (1/initial) = 13; N (2) = 15; N (3) = 15; N (4) = 15; N(5) = 15; N (6/final) = 15

At least at the phrase-level, the shape of the overall movement can be taken as good evidence for phrasing in the domain of something like an intonational phrase – i.e. as evidence that what has here been labeled "phrase" indeed largely

captures a phrase in the phonological sense. If the findings of a similar intonational shape across the word or a unit approximating it in size turned out to be robust, this would count as evidence for a lower-level type phrase as well. Whether this lower-level domain corresponds to the word or rather a small phrase above the morphosyntactic word such as an extended NP with preceding adjectives or a PP with embedded Noun, is another question.

4.3.2 Duration

Our variety of Quechua uses vowel length to distinguish meanings both at the level of lexemes and that of grammatical suffixes, for example *wata* 'year' vs. *wa:ta* 'domestic animal" vs. *wa:ta:* (keep.animals.1s) 'I keep it [i.e. an animal]' (cf. Parker 1976: 51). As can be seen from these examples and from the literature, the positioning and multiple occurrence of such lengthened syllables is not constrained at the word level. However, when a syllable that has a long nucleus and is open is combined with a suffix beginning with two consonants, the first of those becomes the coda of the first syllable, and the vowel is said to be shortened (CV: + CCV -> CVC.CV, with the exceptions of nominal roots ending in a long vowel and the vowel lengthening used to encode verbal first person) in order to conform to a maximal syllable structure of CVC or CV: (cf. Parker 1976: 51–52). Nonetheless, as already mentioned, in our data we find a large number of instances of severely reduced word- or phrase-final syllables, which effectively yields spoken "super heavy" final syllables with structures like CVCC or CV:CC from a combination of the penult with this reduced final syllable. Hence it seems that often, the difference between long and short vowels manifests as unreduced but short vs. fully elided vowels. Apart from contravening proposed constraints on the syllable structure of our variety of Quechua, this process of course also has the effect of shifting syllable position from the right word boundary one step to the right (the "original" penult with the reduced final syllable becoming the "new" final syllable). In this situation, our expectation is not to find a straightforward encoding of a fixed prominent position at the word level by means of duration. Indeed, the results of plotting the left-right-coefficient for duration against syllable position from the right word boundary are very similar to those of the overall coefficient already discussed: basically no large outstanding differences between the positions, especially between penult and ultima (see Figure 2). This changes again when only considering words of length 4, but more interestingly, syllable structure has a much stronger effect: Figure 6 compares the same measurements restricted to syllables of the form CV (orange) to those of the form CVC (purple).

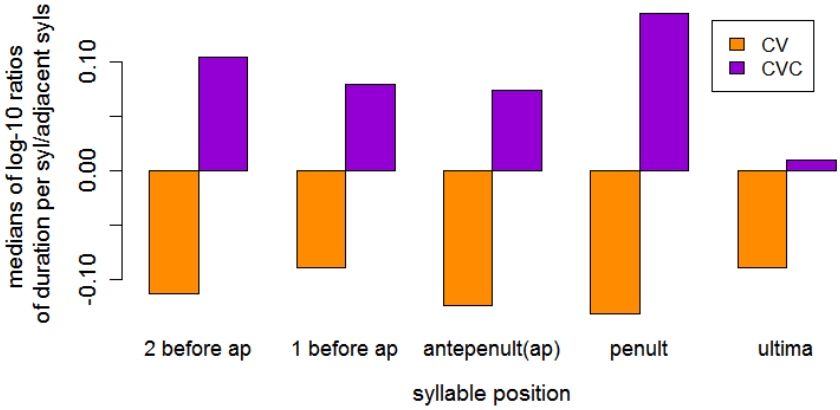

Figure 6: Barplots comparing medians of log-10 ratios of the left-right coefficient for duration between CV-syllables (orange) and CVC-syllables (purple), ordered according to syllable position in the word from the right word boundary. For CV-syllables: N (2 before antepenult) = 7; N (1 before antepenult) = 28; N (antepenult) = 67; N (penult) = 126; N (ultima) = 119. For CVC-syllables: N (2 before antepenult) = 4; N (1 before antepenult) = 7; N (antepenult) = 41; N (penult) = 84; N (ultima) = 79

Penults of the form CVC seem to be much longer than their surrounding syllables in comparison with penults of the form CV. Note also that the scale of the y-axis here indicates much larger differences than e.g. in the overall results (Figure 2). Consider here that we cannot tell whether the adjacent syllables for each individual syllable of this form also had the same form, thus it is possible that all of the penultimate CVC syllables were surrounded by syllables of shorter structure. If that were the case however, it would suggest a distribution of CVC-syllables sensitive to syllable position within the word, which would contradict Parker's description and would be interesting in itself. However, it would also be the case that a large share of the syllables in word-final position annotated here as CVC are the product of final syllable reduction and are thus "former" penults. Thus, if a process existed to produce penults of the form CVC, it would work against that one producing final syllables of that form. In fact, many of the CVC-penults might actually have ended up as final syllables in this data, but the

relative length distinction persists. It is thus plausible that syllables in penultimate position receive some kind of prominence that is realized at least partially through duration, but that this is often obscured by the length difference realized on behalf of syllable structure. As with all findings presented here, this one also needs to be supported by further analysis of a larger dataset.

4.3.3 Intensity

Intensity is usually not a very good correlate of metrical prominence in most languages, although it was once thought to be the main correlate in languages with word-level stress (Beckman 1986). Here, it presents an interesting addition to the results so far: From Figure 2, we can see that of all measures of realizational strength we have looked at so far, intensity (green) shows the penult to stand out from the other positions most clearly without the application of any further conditions. However, looking at the scale for the y-axis in Figure 2, we can see that none of the parameters there reach a median value above 0.04 (corresponding to a ratio of 1.096478 to 1), so also the strength of the effect of intensity here should not be overestimated.

4.4 Information structural categories

In many proposals, information structural categories play an important role for the assignment of prominence at a high level of prosodic structure. Especially, the category of focus is often associated with a particular pitch accent (Büring 2012, among many others) and many language specific annotation systems (ToBIs) implementing the autosegmental-metrical model of intonation (Pierrehumbert 1980) include different intonational categories for different kinds of foci such as information focus and contrastive focus, including the ToBI for peninsular Spanish Sp_ToBI (Beckman et al. 2002; Estebas-Vilaplana & Prieto 2008). An additional assumption that has been proven empirically for many languages is that such high-level prominence markings of foci occur at the same site as other word- or phrase-level prominences, i.e. that a focused word will receive a particular accentual prominence on the syllable that is metrically already most prominent. Several recent proposals have cast the universality of this claim into doubt (cf. Kügler & Genzel 2012; Féry 2013). It seems to be the case that at least in some languages, the hierarchy of focus strength proposed in Féry (2013: 688–690), going from broad information focus to narrow corrective focus necessitates a realization via marked prosody only at higher points of the scale. It is nevertheless promising to look for loci of prominence via information structural categories. It comes as somewhat of a surprise, then, that no clear overall results are ob-

tained when applying this difference in labeling. Reasons for this might lie in the relatively broad labeling decision regarding focus (see §3.3.2), which might have to be refined for further studies, adopting Féry's hierarchy of focus strength, and which could allow for too much variation, or in the sample size. The same holds for the difference between those syllables labelled as "given", and those labelled as "new", and again speculation about reasons for this will lead us most immediately to categories labelled too broadly and the small sample size. However, a distinction that does yield interesting results is the one between topic and comment. See Figure 7 for a comparison of the values for topic (dark green) and comment (blue). Why this is the one information structural category yielding mentionable results (again favoring the penultima), is at this stage open to speculation. A possible factor might be that topics that are fully realized are often contrastive topics or topic shifts and hence focal in the sense that they evoke salient alternatives; the way the data was annotated has a bias towards labeling parts of utterances of comment insofar as that utterances without realized topics were labelled as "comment" (and not e.g. as "thetic"), so "topic" might actually often label a subset of focused parts of speech, namely those focused contrastively. Contrastive foci rank relatively high in Féry's hierarchy of focus strength and are thus more likely to be realized with marked prosody in her account.

A related explanation might be that realized topics are often only nominal constituents, again making them a more "narrow" label. Both these explanations receive incidental support by the fact that there is a total of 466 syllables labelled "comment" in the sample used versus only 235 labelled "topic". Further investigation into this is again needed.

5 Discussion – what we've got, what needs to be done and what's feasible

The results we have obtained so far are promising in that they demonstrate that it is indeed possible with this method to say something about consistent positions of realizational strength in a language where phonological prominence patterns are unknown, thus providing an empirical basis for hypotheses about these positions and the processes affecting their realization. The most convincing result for the utterances analyzed in this pilot corpus is that of prominence on the penult of the (prosodic) word, as demonstrated by the overall coefficient for 4-syllable words and under the condition topic and pitch range; duration when controlled for syllable structure and intensity overall. A second result is that of a possible

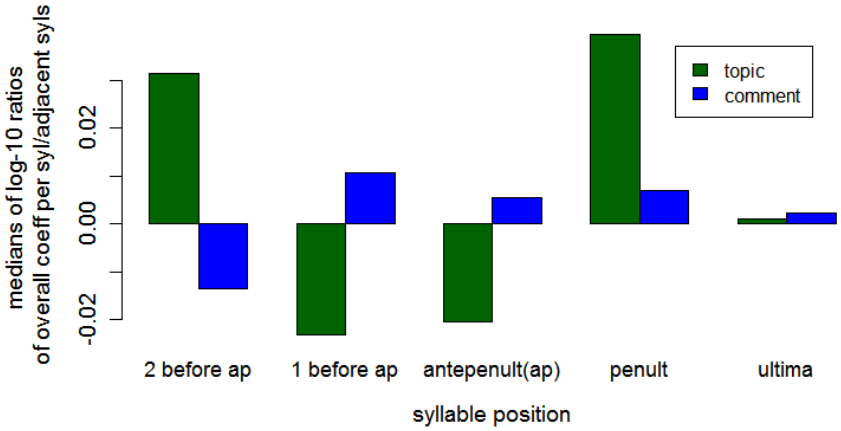

Figure 7: Barplots comparing medians of log-10 ratios of the overall left-right coefficient between syllables labeled "topic" (dark green) and "comment" (blue), ordered according to syllable position in the word from the right word boundary. For "topic": N (2 before antepenult) = 5; N (1 before antepenult) = 15; N (antepenult) = 37; N (penult) = 93; N (ultima) = 84. For "comment": N (2 before antepenult) = 6; N (1 before antepenult) = 33; N (antepenult) = 97; N (penult) = 180; N (ultima) = 194

secondary prominence on the initial syllable, as indicated by the overall and pitch range coefficient on 4-syllable words, and that for duration of CVC syllables. The most general caution regarding any of these hypotheses concerns the size and lack of balance of the corpus, and the consequent lack of any inferential statistics corroborating the general applicability of the results' predictions. We intend to improve upon this state of affairs with a follow-up study on a larger sample. More specific limitations are discussed in the following. It would be useful to test the method as explained here on a comparable corpus of a language where prominence positions are well known, such as peninsular Spanish.[10] This is also something we intend to remedy in the future. Comparable research, reported on in Pamies Bertrán (1994 and especially 1996), has also used a coefficient composed of ratios of one acoustic parameter for one syllable by that of its neighbors to study acoustic realization of prominent positions in several languages where these positions are known. Although the calculations used to derive at his results,

[10] We thank Paolo Roseano for this and several other very useful suggestions on how to improve this work.

described in Pamies Bertrán (1996: 27–29), remain somewhat vague, the results themselves are similarly promising in that they suggest a general realization of prominences by a combination of these cues and a compensatory mechanism between them that is at work to differing degrees in the different languages under his discussion. His sample size for each of the seven languages analyzed (Catalan, English, French, Italian, Portuguese (lumping European and Brazilian Portuguese together, which we find questionable from a prosodic point of view), Russian and Spanish) is even smaller than ours for Quechua, hence the same general limitations apply. However, one very consistent result of his (reported on at greater length for Spanish in Pamies Bertrán 1994) is that prominent positions adjacent to each other ("acentos contiguous" in his terminology, "stress clash" more generally) do not allow for a consistently strong realization by any of the parameters. This has a possible bearing on our results as well, since it might help explain why no good results could be obtained in almost any condition for 2- and 3-syllable words. This would follow immediately if both the penult and the initial syllable were indeed prominent, creating clashes to be resolved by the phonetic realization (see also Hintz 2006 for a similar observation on a central Quechua variety very close to ours). It would especially be the case if what is seen as realizationally strong in the initial is mainly a rise in F0, whereas being a combination of a fall in F0 plus intensity and optional durational prominence in the penult; the rises and falls would conflict in 2-syllable words and create movements that cannot be disentangled by this method alone. This is one of the more specific limitations of this method, i.e. that in the domain of intonation it cannot properly differentiate between such phenomena as late and early peaks. The comparison of the coefficients for mean F0 and pitch range does allow for some more fine-grained intonational profiling of words of a given length, but it cannot resolve the issue of consistently late peaks versus peaks within the accented syllable by itself. Another related issue is that of reduction and prosodic phrasing. While in theory it would be possible to calculate mean constellations of F1–F2 for each intended vowel type in the corpus and then calculate their reduction by centralization by measurement of Euclidean distances from that mean and their durational reduction for each individual vowel token, in practice this would mean individually checking each vowel due to issues in automatic formant measurement in Praat, greatly reducing the advantage of automatization this method aims at. Hence, the clearly existing reduction processes in our variety of Quechua are not very well captured by this method. However, they are prime candidates as indicators for prosodic boundaries and hence very important to our general endeavor of determining prominence positions. To sum up and refine our desires for further

investigations, they should therefore consist of acoustic measurements of the kind described here on a larger corpus of utterances where reductions are annotated as prosodic boundaries, preferably on ones of syllable length 4 and greater in order to better disentangle primary and secondary prominences and their differing realizations.

Acknowledgments

The research presented here has been conducted as part of the DFG-funded project *Zweisprachige Prosodie: Rhythmus, Metrik und Intonation zwischen Spanisch und Quechua* at FU Berlin. As part of this project, we undertook a fieldtrip to Huari, Conchucos, Peru in the summer of 2015, from which the data used in this research stems. We very gratefully acknowledge the support of the Deutsche Forschungsgemeinschaft and would also like to thank the speakers we recorded for their time and willingness to participate in our experiments.

References

Abaurre, Maria B. & Charlotte Galves. 1998. As diferenças rítmicas entre o Português Europeu e o Português Brasileiro: uma abordagem otimalista e minimalista. The rhythmic differences beween European Portuguese and Brazilian Portuguese. *Documentação de Estudos em Lingüística Teórica e Aplicada* 14(2). 377–403.

Adelaar, Willem F. H. 1977. *Tarma Quechua: Grammar, Texts, Dictionary*. Lisse: Peter de Ridder.

Anderson, Anne H., Miles Bader, Ellen Gurman Bard, Elizabeth Boyle, Gwyneth M. Doherty, Simon Garrod, Stephen Isard, Jacqueline Kowtko, Jan McAllister, Jim Miller, Catherine Sotillo, Henry S. Thompson & Regina Weinert. 1991. The HCRC Map Task Corpus. *Language and Speech* 34. 351–366.

Auer, Peter. 1993. *Is a rhythm-based typology possible?: A study of the role of prosody in phonological typology* (KontRI Working Paper 21). Konstanz: Universität Konstanz.

Auer, Peter. 2001. Silben- und akzentzählende Sprachen. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals // Sprachtypologie und sprachliche Universalien // La typologie des langues et les universaux linguistiques: An international handbook // Ein internationales Handbuch // Manuel international*, vol. 2, 1391–1399. Berlin, New York: de Gruyter.

Aurrekoetxea, Gotzon, Iñaki Gaminde, Leire Gandarias & Aitor Iglesias. 2012. Prosodic Variation in the Basque Language: Stress Areas. In Xosé A. álvarez Pérez (ed.), *Proceedings of the International Symposium on Limits and Areas in Dialectology (LimiAr), Lisbon, 2011.* Lisboa.

Austin, John L. 1962. *How to Do Things with Words: The William James Lectures delivered at Harvard University in 1955.* Oxford: Oxford University Press.

Baumann, Stefan. 2006. *The Intonation of Givenness: Evidence from German.* Tübingen: Niemeyer.

Becker, Thomas. 1996. Zur Repräsentation der Vokallänge in der deutschen Standardsprache. *Zeitschrift für Sprachwissenschaft* 15(1). 3–21.

Beckman, Mary E. 1986. *Stress and non-stress accent* (Netherlands Phonetic Archives 7). Dordrecht: Foris.

Beckman, Mary E., Manuel Díaz-Campos, Julia Tevis McGory & Terrell A. Morgan. 2002. Intonation across Spanish, in the Tones and Break Indices framework. *Probus. International Journal of Latin and Romance Linguistics* 14(1). 9–36.

Boersma, Paul & David Weenink. 2017. *Praat: Doing phonetics by computer [Computer program].* Version 6.0.30. http://www.praat.org/.

Bolinger, Dwight. 1989. *Intonation and its uses: Melody in grammar and discourse.* London [etc.]: Edward Arnold.

Büring, Daniel. 2012. Focus and Intonation. In Gillian Russell & Delia G. Fara (eds.), *Routledge companion to philosophy of language (Routledge Philosophy Companions)*, 103–115. New York, NY: Routledge.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles N. Li (ed.), *Subject and topic*, 27–55. New York: Academic Press.

Dauer, Rebecca. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11. 51–62.

Delforge, Ann M. 2008. Unstressed Vowel Reduction in Andean Spanish. In Laura Colantoni & Jeffrey Steele (eds.), *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, 107–124. Somerville, MA: Cascadilla Proceedings Project.

Delforge, Ann M. 2011. Vowel Devoicing in Cusco Collao Quechua. In Wai-Sum Lee & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, 556–559. Hong Kong: City University of Hong Kong.

Dufter, Andreas. 2003. *Typen sprachrhythmischer Konturbildung* (Linguistische Arbeiten 475). Tübingen: Niemeyer.

Estebas-Vilaplana, Eva & Pilar Prieto. 2008. La notación prosódica del español: una revisión del Sp_ToBI. *Estudios de fonética experimental* 17. 263–283.

Féry, Caroline. 2013. Focus as prosodic alignment. *Natural Language & Linguistic Theory* 31(3). 683–734.

Gómez Rendón, Jorge. 2006. Interpersonal aspects of evidentiality in Ecuadorian Quechua. In Silke Hamann & Roland Pfau (eds.), *ACLC Working Papers*, vol. 1, 37–50. Amsterdam: ACLC.

Grice, Martine, D. Robert Ladd & Amalia Arvaniti. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17(2). 143–185.

Grice, Martine & Michelina Savino. 1997. Can pitch accent type convey information status in yes-no questions? In Kai Alter, Hannes Pirker & Wolfgang Finkler (eds.), *Concept to Speech Generation Systems: Proceedings of a Workshop in conjunction with 35th Annual Meeting of the Association for Computational Linguistics*, 29–38. Madrid: Universidad Nacional de Educación a Distancia.

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Hintz, Daniel. 2000. *Caracteristicas distintivas del quechua de Corongo*. Lima: Instituto Lingüistico de Verano.

Hintz, Diane M. 2006. Stress in South Conchucos Quechua: A Phonetic and Phonological Study. *International Journal of American Linguistics* 72(4). 477–521.

Hualde, José Ignacio, Gorka Elordieta, Iñaki Gaminde & Rajka Smiljanic. 2008. From pitch accent to stress accent in Basque. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7 (Phonology and phonetics 4-1)*, 547–584. Berlin, New York: Mouton de Gruyter.

Hyman, Larry M. 2009. How (not) to do phonological typology: The case of pitch-accent. *Language Sciences* 31(2–3).

Hyman, Larry M. 2014. Do All Languages Have Word Accent? In Harry van der Hulst (ed.), *Word stress: Theoretical and typological issues*, 56–82. Cambridge: Cambridge University Press.

Ka, Omar. 1989. Wolof Syllable Structure: Evidence from a Secret Code. In Joyce Powers & Kenneth d. Jong (eds.), *Proceedings of the Fifth Eastern States Conference on Linguistics*, 261–274. Columbus, Ohio.

Ka, Omar. 1994. *Wolof Phonology and Morphology*. Lanham, Maryland: University Press of America.

Krifka, Manfred. 2007. Basic Notions of Information Structure. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *Interdisciplinary Studies on Information Structure: The notions of information structure* (Working Papers of the

SFB 632 Interdisciplinary Studies on Information Structure 6), 13–56. Potsdam: Potsdam: Universitäts-Verlag.
Kügler, Frank & Susanne Genzel. 2012. On the Prosodic Expression of Pragmatic Prominence: The Case of Pitch Register Lowering in Akan. *Language and Speech* 55(3). 331–359.
Liberman, Mark & Alan S. Prince. 1977. On Stress and Linguistic Rhythm. *Linguistic Inquiry* 8(2). 249–336.
Muntendam, Antje G. 2010. *Linguistic Transfer in Andean Spanish: Syntax or Pragmatics?* Urbana-Champaign: University of Illinois at Urbana-Champaign dissertation.
Muntendam, Antje G. 2012a. Information structure and intonation in Andean Spanish. In *LSA Annual Meeting Extended Abstracts 3*.
Muntendam, Antje G. 2012b. *Bilingual intonation: a case study of Andean Spanish*. Bangor: Wales.
Muntendam, Antje G. & Francisco Torreira. 2016. Focus and prosody in Spanish and Quechua: Insights from an interactive task. In Meghan E. Armstrong, Nicholas Henriksen & Maria d. M. Vanrell (eds.), *Intonational Grammar in Ibero-Romance: Approaches across linguistic subfields*, 69–89. Amsterdam: John Benjamins.
Muysken, Pieter. 1995. Focus in Quechua. In Katalin é Kiss (ed.), *Discourse configurational languages*, 375–393. Oxford: Oxford University Press.
Muysken, Pieter & Antje G. Muntendam. 2016. Interfacing interfaces: Quechua and Spanish in the Andes. In Susann Fischer & Christoph Gabriel (eds.), *Manual of Grammatical Interfaces in Romance*, 607–633. Berlin, Boston: de Gruyter Mouton.
O'Rourke, Erin. 2005. *Intonation and Language Contact: A Case Study of two Varieties of Peruvian Spanish*. Urbana-Champaign: University of Illinois dissertation.
O'Rourke, Erin. 2007. Intonation in Quechua: Questions and Analysis. In Jürgen Trouvain & William J. Barry (eds.), *ICPhS XVI, Saarbrücken Germany, 6 - 10 August 2007: 16th international congress of phonetic sciences*. Saarbrücken.
O'Rourke, Erin. 2008. Correlating Speech Rhythm in Spanish: Evidence from Two Peruvian Dialects. In Joyce Bruhn de Garavito & Elena Valenzuela (eds.), *Selected proceedings of the 10th Hispanic Linguistics Symposium*, 276–287. University of Western Ontario.
O'Rourke, Erin. 2009. Phonetics and phonology of Cuzco Quechua declarative intonation: An instrumental analysis. *Journal of the International Phonetic Association* 39(3). 291–312.

O'Rourke, Erin. 2010. Dialect Differences and the Bilingual Vowel Space in Peruvian Spanish. In Marta Ortega-Llebaria (ed.), *Selected proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, 20–30. Somerville, MA.

O'Rourke, Erin. 2012. The realization of contrastive focus in Peruvian Spanish intonation. *Lingua* 122(5). 494–510.

Pamies Bertrán, Antonio. 1994. Los acentos contiguos en español. *Estudios de fonética experimental* 6. 91–111.

Pamies Bertrán, Antonio. 1996. Consideraciones sobre la marca acústica del acento fonológico. *Estudios de fonética experimental* 8. 11–49.

Parker, Gary. 1976. *Gramática quechua: Ancash-Huailas*. Lima: Ministerio de Educación.

Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation*. Bloomington: MIT dissertation.

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

Reich, Uli. 2002. *Freie Pronomina, Verbalklitika und Nullobjekte im Spielraum diskursiver Variation des Portugiesischen in São Paulo* (Romanica Monacensia 62). Tübingen: Narr.

Reich, Uli & Martin Rohrmeier. 2014. Batidas latinas: On rhythm and meter in Spanish and Portuguese and other forms of music. In Javier Caro Reina & Renata Szczepaniak (eds.), *Syllable and word languages*, 391–420. Berlin: de Gruyter.

Rialland, Annie & Stéphane Robert. 2001. The Intonational System of Wolof. *Linguistics* 39(5). 893–939.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116.

Searle, John R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Trager, George L. 1945. Analysis of a Kechuan Text. *International Journal of American Linguistics* 11(2). 86–96.

Trubeckoj, Nikolaj S. 1939. *Grundzüge der Phonologie: Publié avec l'appui du Cercle linguistique de Copenhague et du ministère de l'instruction publique de la République tchéco-slovaque*. Vol. 7 (Travaux du Cercle Linguistique de Prague). Prag.

Vallduví, Enric. 1992. *The informational component*. New York: Garland.

van der Hulst, Harry. 1999. Word accent. In Harry van der Hulst (ed.), *Word prosodic systems in the languages of Europe*, vol. 20 (Empirical approaches to language typology 4), 3–115. Berlin, New York: Mouton de Gruyter.

van der Hulst, Harry. 2010. Word accent systems in the languages of Europe. In Rob Goedemans, Harry van der Hulst & Ellen van Zanten (eds.), *A survey of word accentual patterns in the languages of the world*, 429–507. Berlin: de Gruyter.

van Rijswijk, Remy & Antje G. Muntendam. 2014. The prosody of focus in the Spanish of Quechua-Spanish bilinguals: A case study on noun phrases. *International Journal of Bilingualism* 18(6). 614–632.

Voisin-Nouguier, Sylvie. 2002. *Relations entre fonctions syntaxiques et fonctions sémantiques en wolof*. Lyon: Université Lumière Lyon dissertation.

Weber, David J. 1986. Information Perspective, Profile and Patterns in Quechua. In Wallace L. Chafe & Johanna Nichols (eds.), *Evidentiality: The linguistic coding of epistemology* (Advances in discourse processes 20), 137–155. Norwood, N.J.: Ablex publishing corporation.

Wetzels, Leo & Sérgio Meira. 2010. A Survey of South American stress systems. In Rob Goedemans, Harry van der Hulst & Ellen van Zanten (eds.), *A survey of word accentual patterns in the languages of the world*, 313–380. Berlin: de Gruyter.

Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann & Han Sloetjes. 2006. ELAN: A Professional Framework for Multimodality Research. In Nicoletta Calzolari, Khalid Choukri, Aldo Gangemi, Bente Maegaard, Joseph Mariani, Jan Odijk & Daniel Tapias (eds.), *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'06)*, 1556–1559. Genoa, Italy.

Wölck, Wolfgang. 1972. *Especificación y foco en quechua*. San Marcos.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.