STATISTICAL CO-OCCURRENCE RESTRICTIONS IN OROMO CONSONANTS [*]

Avery Ozburn
University of Toronto

1. Introduction

While most studied cases of consonant harmony are categorical, some languages show gradient harmony tendencies that are similar to patterns that are categorical in other languages (see e.g. Arsenault 2012, Brown 2008). This paper presents a new case of statistical consonant harmony in the eastern (Harar) dialect of Oromo (Cushitic, Ethiopia). In Oromo, there is a clear bias towards laryngeal agreement in stops, both in ejectivity and in voicing; however, these restrictions are not absolute, and there are many cases without agreement. Further, a statistical regressive directionality asymmetry appears within the overall statistical agreement pattern. This non-categorical directionality trend mirrors patterns that are categorical in other languages and gives additional motivation to investigate how phonological theories can account for statistical tendencies. However, the primary goal of this paper is not to give a theoretical account, but rather to report on the preliminary results of a study of laryngeal harmony in Oromo.

The paper is organized as follows: Section 2 provides background on laryngeal harmony and on Oromo, Section 3 looks at the Oromo data, Section 4 gives discussion of the results, and Section 5 concludes.

2. Background

2.1 Laryngeal harmony

Among languages that have laryngeal harmony, there are several generalizations that will help to situate the data in the remainder of this paper. This section outlines the patterns observed in known cases of voicing and ejective harmony across a variety of languages. Since the majority of reported cases of laryngeal harmony deal with examples that are categorical, this section focuses on exceptionless or near-exceptionless cases of laryngeal harmony, looking at similarity (2.1.1), locality (2.1.2), directionality (2.1.3), and alternations (2.1.4), all of which are important properties of harmony.

2.1.1 Similarity

Similarity is often a crucial element of consonant harmony systems (e.g. Rose and Walker 2004). This similarity can manifest in several ways, as discussed in more detail in this section. I focus on place of articulation and laryngeal features, as they are the features of relevance in considering harmony in Oromo.

[*] Thank you to my Oromo consultant, Tamam Youssouf, for providing the data in this paper. Thank you as well to the audience of the Agreement by Correspondence Conference held at the University of California Berkeley in May 2014, to the University of Toronto field methods class, and to my supervisors Keren Rice and Peter Jurgec, for helpful discussion about this research.

Actes du congrès annuel de l'Association canadienne de linguistique 2014.
Proceedings of the 2014 annual conference of the Canadian Linguistic Association.

Place of articulation often has a significant effect on laryngeal co-occurrence restrictions. In particular, stops agreeing in place of articulation are more likely to be required to agree in laryngeal features than stops that disagree in place (e.g. Hansson 2001, 2010). Similarly, in languages with restrictions on the number of ejectives in a word, two ejectives are more likely to be allowed in a word if they agree in place (Gallagher 2013). Indeed, in a large proportion of languages with laryngeal restrictions, stops agreeing in place behave specially (Hansson 2001, 2010). For example, in Chol (Mayan), pairs of identical ejectives are allowed, while pairs of non-identical ejectives are banned (Gallagher 2010). Similar kinds of homorganicity restrictions apply in a number of cases of voicing harmony (see Hansson 2001, 2010 for examples). Thus, it is important to ask whether generalizations differ between homorganic and heterorganic stops for both types of laryngeal harmony in Oromo.

In Oromo, ejective and voicing harmony are of particular interest. In terms of ejective harmony, the most common form of ejective restrictions cross-linguistically is between ejectives and plain voiceless stops. In other words, if a language has voiced stops, they will often not participate in any co-occurrence restrictions with ejectives; stops must be similar in voicing in order to participate in such restrictions (Hansson 2001, 2010). For example, Hausa has harmony between voiceless pulmonics and ejectives as well as between voiced pulmonics and implosives, but not between voiced pulmonics and ejectives (Hansson 2001, 2010). Thus, this paper looks solely at restrictions that exist between ejectives and plain voiceless stops in Oromo, and does not consider [constricted glottis] (a feature I use to distinguish ejective and implosive consonants from plain voiceless and voiced consonants) restrictions between stops that differ in voicing.

2.1.2 Locality

Consonant harmony, by definition, only applies at a distance, across a vowel; strictly adjacent consonants affecting each other is generally not considered to be part of the scope of harmony (see, for instance, Rose and Walker 2004 and Hansson 2001, 2010). In terms of locality, stops in adjacent syllables are more likely to be subject to co-occurrence restrictions than stops at a greater distance. This property is common to all types of consonant harmony, not just laryngeal harmony. Indeed, typologically, it is possible for languages to have consonants interacting only across a vowel but not at a greater distance, but the reverse is not found (e.g. McMullin and Hansson 2013). This property justifies the decision in this paper to restrict the database to words with stops in adjacent syllables; more distant stops often behave differently, and also are more likely to be separated by morpheme boundaries.

2.1.3 Directionality

In terms of directionality, laryngeal harmony is often a regressive process triggered by the marked feature value (Hansson 2001, 2010). For example, in Ngizim (Chadic), sequences of a voiceless stop followed later in the word by a voiced stop, like T…D, are forbidden, while the reverse co-occurrence, namely D…T, is allowed (Hansson 2001, 2010). The same holds of ejective harmony. For instance, in Souletin Basque and Bolivian Quechua, a word with an ejective followed later by a voiceless stop is allowed, but the reverse is not (Gallagher 2010). This property is relevant to directionality; in these languages, [+voice] or [+cg] consonants trigger leftward harmony, but do not trigger

harmony rightwards. This paper examines similar kinds of directionality restrictions in Oromo. However, in Oromo, unlike in previously studied languages like Ngizim, these directionality restrictions are not categorical, but rather statistical. We will see in more detail below that within morphemes with homorganic stops, there are many more Oromo words with a voiced stop followed by a voiceless stop in the following syllable than the reverse order, and that this difference approaches significance.

2.1.4 Alternations

Finally, most cases of laryngeal harmony discussed in the literature are morpheme structure constraints, and therefore do not involve alternations (Hansson 2001). Of all languages with voicing harmony, only Berber and possibly Kera and Yabem are known to have cross-morpheme non-local voicing alternations (Hansson 2001, 2004, 2010); even fewer languages have ejective alternations, with Lezgian being the only known case (Ozburn and Kochetov 2013). Thus, laryngeal harmony tends to be simply a pattern within the lexicon. In Oromo, there do not appear to be non-local laryngeal alternations; I will examine only morpheme-internal patterns.

2.2 Oromo

2.2.1 Background about Oromo

The language of focus in this paper is Oromo (Cushitic). It is spoken as a native language by approximately 17.5 million people, mostly in the Oromia region of Ethiopia, but also in parts of Kenya and other countries (Lewis et al. 2014). The data in this paper is from the Harar (eastern) dialect, which was the native language of the consultant for this paper. However, preliminary work on a dictionary based Oromo as spoken in southwestern Ethiopia suggests that this dialect also has very similar restrictions on disagreeing laryngeal combinations (see Ozburn 2014).

As an Ethiopian language, Oromo has been in extensive contact with a number of other languages, particularly (Ethio-)Semitic languages such as Arabic and Amharic; many native speakers of Oromo, particularly those in the older generation, are also fluent speakers of Arabic and Amharic from the school system (Youssouf p.c.). Such loans are important to know about because they could behave differently with respect to phonological processes (see e.g. Kang 2011 for a recent overview). For example, Ethio-Semitic languages like Amharic and Chaha have their own laryngeal co-occurrence restrictions (Rose and King 2007), so the Amharic loans in Oromo could change results or even make co-occurrence restrictions appear where none exist. For this paper, the consultant was fluent in Amharic and Arabic, and was therefore able to indicate when words were Semitic loans. Such words were excluded from the database.

2.2.2 Consonant inventory of Oromo

The part of the Oromo inventory relevant to this research is the stop/affricate inventory. Based on work with the consultant as well as confirmation from an Oromo-English dictionary (Gamta 1989), the stop/affricate inventory of Oromo is as follows:

(1) Oromo stop/affricate inventory

		Bilabial	*Coronal*[1]	*Post-Alveolar*	*Palatal*	*Velar*	*Glottal*
Stops and Affricates	*Voiceless*		**t, tt**	**ttʃ**		**k, kk**	ʔ
	Voiced	**b, bb**	**d, dd**	**dʒ, ddʒ**		**g, gg**	
	Ejective	**p', pp'**	**t', tt'**	**tʃ', ttʃ'**		**k', kk'**	
	Implosive		**ɗ, ɗɗ**				

In this paper, we are particularly interested in the bolded consonants, namely the non-glottal stops and affricates[2]. From this table, we can see that Oromo has laryngeal contrasts in stops that vary from a two-way contrast for bilabials (b, p'), to a three-way contrast for post-alveolars (tʃ, dʒ, tʃ') and velars (k, g, k'), to a four-way contrast for coronals (t, d, t', ɗ). There are two types of voiceless consonants, plain voiceless and ejectives, and two types of voiced ones, plain voiced and implosives. Further, there are two types of constricted glottis consonants, ejective and implosive, and two types of plain consonants, voiced and voiceless.

Note that there are several gaps in the inventory, as well as gemination contrasts in many consonants[3]. The gaps are crucial to a more detailed study of the language and to an understanding of what data was used for this study; they will be considered further in the next section. Gaps outside of the stops are also relevant to understand the scope of this paper. Within the fricatives and sonorants, there are no laryngeal contrasts, with the only exception being a voicing contrast within the coronal fricatives (Gamta 1989). However, the consultant indicated that /z/ is a marginal sound in Oromo, occurring primarily in loan words and often adapted to /s/ by older speakers. Due to this lack of contrast in other manners, it is reasonable to limit the study to stops and affricates, which show multiple robust laryngeal distinctions.

2.2.3 Phonotactic constraints

In order to understand some aspects of the data, it is crucial to note several aspects of Oromo phonotactics and morphology. This section will consider the behaviour of /p'/, the distribution of geminates, possible morpheme-internal clusters, word length, and affixation properties.

First, while the ejective bilabial stop /p'/ exists in Oromo, it is rare and has limited distribution compared to other stops. The consultant strongly dispreferred /p'/ in initial position and did not accept any nonsense words containing it in that position. This sound is also relatively limited in word-medial position compared to other consonants, with far fewer examples than for other stops. For that reason, combined with the fact that the

[1] While post-alveolar affricates are also coronal, I will use the term 'coronal' to refer only to the consonants in this column. In at least some dialects of Oromo, the voiceless, voiced, and ejective coronal stops are dental, while the implosive is retroflex (Owens 1985). The term 'coronal' best captures this class.

[2] For the remainder of this paper, the word 'stops' will be understood to include both stops and affricates, which pattern similarly and will be considered together.

[3] Geminates are included in the table above because not all consonants have both non-geminate and geminate versions. This fact is particularly important for the post-alveolar affricates.

voiceless bilabial does not exist in the language, bilabials were excluded from the current data and are left to future research.

Second, geminates are possible only intervocalically; it is impossible in Oromo to have geminates word-initially, word-finally, or next to another consonant. This fact is confirmed by the lack of geminates anywhere other than intervocalically in the dictionary, as well as by the intuitions of the native speaker consultant. Thus, in the CV(C)CV words to be studied in this paper, geminates are possible only as the second consonant, because they cannot be initial. This fact becomes crucial for the post-alveolars, since the plain post-alveolar affricate can occur only as a geminate. As such, due to the restrictions on geminates in the language, the plain post-alveolar affricate cannot occur as C1. As with the bilabials, the post-alveolars will therefore be excluded from the present study.

Third, Oromo has very limited morpheme-internal consonant clusters, and the only stop-stop clusters that are possible morpheme-internally are geminates[4]. This fact was determined through dictionary work and consultation with the speaker, and it will be important to understanding the database. I assume here that all geminates have a single laryngeal quality. For example, in a word like *t'ikk'o:* 'small or little', the geminate *kk'* is considered ejective. This decision follows the notations in the dictionary of Gamta (1989), as well as the consultant's intuitions.

Fourth, the majority of Oromo morphemes have two syllables; words that are longer are typically multi-morphemic, and there are few monosyllabic roots. For the purposes of this study, only the first two syllables of words were considered. While the second and third syllables would be equally close together in terms of the locality properties discussed above, they were not considered because of the possibility of multi-morphemic forms. Since Oromo has no prefixes besides reduplication (Gamta 1989), the initial two syllables, or at least the initial CVC sequence containing both relevant consonants, should belong to the same morpheme.

Finally, observing non-local laryngeal alternations in Oromo is likely to be quite difficult. While a number of suffixes begin with stops, the stems to which they attach typically end with consonants, and the suffixal stops either assimilate locally in laryngeal quality or a vowel is epenthesized and there is no assimilation (Lloret 1995). For example, in Western and Eastern Oromo, /t'+t/ becomes [t't'][5], while in Southern Oromo, it becomes [t'it][6] (Lloret 1995). Given these local processes, potential alternations are beyond the scope of the present study.

3. Data

3.1 About the database

It is useful before looking at the data to understand which data is included in the results and why. Due to certain properties of Oromo and other properties common to co-occurrence restrictions cross-linguistically, certain types of consonants and words were

[4] Whether geminates count as stop-stop clusters or not is not relevant to this discussion. It is just important that in all cases with two stops next to each other, they form a geminate and agree in laryngeal quality.

[5] This fact was confirmed by the consultant.

[6] Note that the two stops in this phonetic pronunciation are across a morpheme boundary, and these cases are therefore outside the scope of the harmony considered in this paper.

omitted from the study, as described below. Counts for the ejective harmony database were compiled from field methods data plus some words from an Oromo-English dictionary (Gamta 1989) that were checked with the consultant. Counts for the voicing harmony database were compiled solely from field methods data. While many of the words from the ejective database, including those taken from the dictionary, are eligible for inclusion in the voicing database, they were not included due to concerns that only looking in the dictionary for words with voiceless stops might skew the results on the voicing harmony data. All words in the ejective database contain only ejective and plain voiceless stops/affricates, while all words in the voicing database contain only plain voiced and plain voiceless stops/affricates.

First, only coronal and velar consonants were considered. As shown in the consonant chart above, Oromo also has post-alveolar affricates and bilabials, but they were omitted for reasons that were briefly noted above. The plain voiceless post-alveolar affricate occurs only as a geminate, and Oromo does not have a plain voiceless bilabial. These facts about the language could easily skew the results. Indeed, since geminates cannot occur initially, any voiceless affricate occurring initially must be an ejective. This fact could skew results about directionality effects later in the paper, and therefore the affricates are excluded in this preliminary study. In terms of the voiceless bilabial, there are very strong reasons for excluding it in an initial survey. In many languages in which a segment has no contrastive counterpart for the co-occurrence feature, that segment is neutral and can occur with any other consonants (Hansson 2001). Thus, since the bilabial ejective has no plain voiceless counterpart in Oromo, we expect that it might be able to occur freely with both ejective and non-ejective consonants of other places of articulation. Including it in the data could therefore skew the results by providing more examples of ejectives co-occurring with non-ejectives than we might otherwise expect. Moreover, as described in Section 2, the bilabial ejective is not permitted word-initially in Oromo, which further reduces the possible combinations possible with bilabials consonants. Thus, while the bilabial consonants would be interesting to include in future studies, to see how they behave with respect to the restrictions we find, it is best to exclude it when initially determining whether restrictions exist in the language.

Second, any form that was potentially reduplicated was not considered. Oromo has partial reduplication of initial CV(C) of roots to give a meaning expressing a type of repetition. For example, /tʃʾapsu:/ means ‘to break’, while /tʃʾattʃʾapsu:/ means to break into small pieces (Youssouf, p.c.). This reduplication takes the form of a prefix, and it can either copy the entire first CVC of the base or else copy the first CV and then geminate the initial consonant of the base. Thus, all words beginning with CiVjCiCiVj or CiVjCkCiVjCk were omitted, even if they did not have a meaning that seemed reduplicated. The consultant noted that several forms that looked like reduplication but with a non-reduplicated meaning, such as the word /galgalla/ for ‘evening’, were historically reduplicated. While it is quite likely that not all forms that looked like reduplication were in fact synchronically or historically reduplicated, it is safer to exclude all such forms from the data. Indeed, the potential reduplications would simply add more forms to the cases of agreement, since reduplication results in agreeing sequences. We will see that co-occurrences tend significantly towards agreement anyways, and including potentially reduplicated forms would have resulted only in a greater level of significance. Thus, by excluding potentially reduplicated forms, it is easier to make a more solid conclusion about the tendency towards agreement.

Third, as far as morpheme boundaries could be determined, words in which the stops occurred in different morphemes were not considered, and words derived from the same root were considered as one entry in the database. As mentioned above, laryngeal harmony is almost always a morpheme structure condition, limited to single morphemes, and so including multi-morphemic items could blur any trends that exist only within morphemes. Further, considering words derived from the same root could skew the results. Since Oromo has a fairly rich system of derivational morphology, counting all forms derived from the same root would amount to counting single roots many times, which could certainly change the results.

Finally, the word structures considered in this study were words beginning in CVCV, CVCCV, and CVRCV, where CC is a geminate, R is any sonorant, V is a vowel, and the Cs are the consonants of interest. Since Oromo does not allow many consonant clusters, these were all word types in Oromo that fit the criteria of two stops in adjacent syllables without a morpheme boundary. As noted above, it is possible that this choice of word types misses some cases where the two stops are in the second and third syllables of a longer morpheme. However, single morphemes longer than two syllables are not common in Oromo, so it is unlikely that many words were missed. Oromo also does not have very many CVC words, so the omission of CVC words from the data does not substantially change the results.

One factor that was ignored is the spirantization of /k/ to [x] in the dialect spoken by the consultant. In the eastern dialect of Oromo, /k/ spirantizes to [x] intervocalically and initially. The consultant often hypercorrected and produced [k] in these contexts. However, if he were to produce these words naturally, some of the ones with /k/ would not contain two stops, but rather a stop and a velar fricative or two velar fricatives. However, there are several reasons for keeping these words in the analysis, even beyond the fact that excluding them would reduce the dataset significantly. There is both historical and synchronic evidence for suggesting that the [x] is underlyingly a /k/; it surfaces that way in other dialects, the environment for [x] is more predictable than that of [k], and the consultant's intuition is that it is /k/. Since co-occurrence restrictions are lexical, it is reasonable to suggest that they operate on underlying forms, and thus that these words underlyingly fit the criteria for consideration in ejective restrictions. Moreover, statistical tests were run on effects of place of articulation on ejective agreement; velars did not have an effect, suggesting that including these velars does not change the results, despite the spirantization process. Thus, it is justified to include /k/ in the data for this paper.

The resulting databases consisted of 78 items for ejective harmony and 89 items for voicing harmony. All words in the dictionary that fit the criteria for the ejective harmony database, namely all words of the aforementioned structure with two velar or coronal stops that were both voiceless (either plain or ejective), were included. While the counts for this paper are small, work in Ozburn (2014) suggests that similar results also hold over a much larger Oromo database from a different dialect.

3.2 Lack of categorical harmony

Looking at the Oromo databases described above suggests the language has no laryngeal harmony. As seen in (2a), there are examples of all combinations of ejectives and plain voiceless stops. Further, as seen in (2b), there are examples of all combinations of voiced and voiceless stops. Since either lack of ejective/voiceless combinations or lack of

voiced/voiceless combinations is usually required for laryngeal harmony, it appears that Oromo does not have a laryngeal harmony system.

(2) a. Example ejective data

	C2 Ejective	*C2 Voiceless Stop*
C1 Ejective	**k'**a**t'**ale: 'smart, clever'	**k'**o**t**u: 'to farm'
C1 Voiceless Stop	**t**u**k'**u: 'to touch'	**k**u**t**i 'cut (imp)'

(2) b. Example voicing data

	C2 Voiced	*C2 Voiceless*
C1 Voiced	**g**o**g**a: 'animal skin'	**b**o**k**a 'rain'
C1 Voiceless	**k'**o**b**a 'finger'	**tʃ'**i**t**a: 'grass'

Note that, according to the intuitions of the consultant, who is fluent in the languages from which Oromo typically borrows words, the disharmonic words are not loans and are not morphologically complex. This observation suggests we should not discount them as exceptions, since there is no independent reason for these words to be exceptional. Thus, Oromo does not have categorical laryngeal harmony.

3.3 Statistical harmony

However, counting the number of words in each category shows that not all combinations are equally attested. Tables (3a) and (3b) show the counts for each category (represented by N) on the first line of each cell, followed by observed over expected (O/E) values on the second line. O/E values above 1 (given in bold) show over-representation of a category, meaning it happens more often than expected from a random distribution, while an O/E value below 1 shows under-representation.

(3) a. Observed counts and O/E values, ejective co-occurrences

	C2 Ejective	*C2 Plain*
C1 Ejective	N=39 **1.53** cf. *k'at'ale:*	N=6 0.35 cf. *k'otu:*
C1 Plain	N=4 0.22 cf. *tuk'u:*	N=29 **2.00** cf. *kuti*

(3) b. Observed counts and O/E values, voicing co-occurrence

	C2 Voiced	*C2 Voiceless*
C1 Voiced	N=36 **1.46** cf. *goga:*	N=15 0.57 cf. *boka*
C1 Voiceless	N=7 0.38 cf. *k'oba*	N=31 **1.58** cf. *tʃ'ita:*

From these tables, it is clear that agreement is over-represented, with O/E over 1, while disagreement is under-represented. The significance of this difference can be tested with a chi-square test. For the ejective co-occurrence, we obtain the result of $\chi 2(1, N = 78) = 42.77$, $p < 0.0001$; for voicing, we get $\chi 2(1, N = 89) = 23.73$, $p < 0.0001$. Thus, both are highly significant, and the O/E values for agreement are all above 1, showing that Oromo has significant statistical laryngeal harmony in both ejectivity and voicing. However, crucially, as we saw above, this pattern is not categorical.

3.4 Directionality effect

Returning to tables (3a) and (3b), we can also notice a pattern within the disagreeing cases. Voiced-voiceless sequences are more represented than voiceless-voiced (15 versus 7 words), and similarly ejective-plain sequences are more represented than plain ejective (6 versus 4 words). This result suggests a directionality effect like in Ngizim (see section 2.1.3); there is a preference for right-to-left spreading or agreement of the marked feature ([+voice] or [+cg]), so that cases like voiceless-voiced, in which the marked feature has not spread regressively, are dispreferred. The significance of this effect can be tested using a regression model with a dependent variable of [voice] agreement or [cg] agreement respectively. It is not significant for ejective harmony ($p>0.1$), but it approaches significance for voicing harmony ($p=0.0791$). Results from Ozburn (2014), in which directionality effects are significant, suggest that the significance might increase in future work with a larger database. This result suggests that, at least for voicing harmony, we may have a statistical directionality effect within a statistical harmony system.

4. Discussion

The Oromo laryngeal harmony pattern is of particular interest with respect to studies of consonant harmony more generally. As discussed in this section, it shows properties of both similarity and directionality, yet it differs from many other consonant harmony systems that have been studied in that these are statistical in nature. In this section, I address the overall statistical nature of harmony, the directionality patterns, and similarity, then briefly comment on locality and alternations.

4.1 Statistical harmony

In the preceding sections, we observed that Oromo shows significant statistical laryngeal harmony. This result is interesting for a number of reasons. First, given the large number of disagreeing words relative to the total number of words in the database, it is difficult to

analyze this case as categorical with exceptions: one would expect only a very small number of exceptions compared to harmonic words in such a case. Second, the consultant recognized the pattern at some level by rejecting nonsense words with disagreement, suggesting that harmony might be a part of the grammar of speakers despite being a non-categorical lexical pattern. This result suggests that close examination of statistical harmony might help to better understand categorical harmony.

4.2 Directionality

The most intriguing aspect of the study is the directionality effect. The Oromo data appears to be unique among studied cases of directionality in that it shows a near-significant statistical directionality effect within a statistical harmony system. As discussed in Section 2, there is cross-linguistically a regressive directionality bias in many types of consonant harmony, including laryngeal harmony. The case in Oromo suggests that this bias may be deeper and more fundamental than accounts of categorical harmony would suggest, since in Oromo it exists on a purely statistical level. Accounting for the directionality effect in Oromo could therefore prove crucial to understanding directionality effects in other types of laryngeal and consonant harmony, giving insight into how regressive biases arise and how we should account for regressive directionality. The result also raises an important question to drive future research into other languages with statistical harmony, where directionality restrictions have not been reported. If regressive directionality is truly fundamental to laryngeal harmony, then we might expect other languages to show a bias similar to Oromo, and we would not expect any to show the opposite, progressive directionality bias.

4.3 Similarity

This study looked only at stops, and therefore did not consider stricture similarity as a factor on agreement. Other Oromo consonants do not contrast in laryngeal quality, except for a voicing contrast in coronal fricatives, but the consultant said that /z/ is a marginal sound, occurring primarily in loans and changed to /s/ by older speakers. Thus, similarity in stricture would be difficult to analyze in Oromo.

However, other similarity effects in laryngeal harmony include homorganicity. Effects of homorganicity were tested in this study, but were found not to be significant. Further study of a larger database from a different Oromo dialect in Ozburn (2014) has shown homorganicity to be a significant factor, with homorganic consonants much more likely to agree in laryngeal quality. In the future, a larger database for the Harar dialect should be compiled to test whether homorganicity is also a factor here.

4.4 Locality and alternations

The remaining two topics from the background section, namely locality and alternations, cannot be commented on here, because only stops in adjacent syllables within morphemes were included in the database for this paper. The study of locality and alternations in Oromo laryngeal harmony is left to future research.

5. Conclusions

In conclusion, this paper has introduced a new case of laryngeal harmony, in the Ethiopian language Oromo. I have established that both laryngeal harmony and directionality effects are found in Oromo in a way that is not obvious from surface examination of the lexicon. Statistical patterns in Oromo mirror patterns that are categorical in other languages, and particularly interesting is a regressive directionality effect. This result suggests a need to understand statistical patterns in order to make sense of the categorical patterns usually treated by phonology. In particular, this result contributes to the understanding that patterns that might be considered unmarked can show up only as statistical effects in the lexicon, not categorically and not reinforced by alternations. Such processes could be quite helpful in understanding broader harmony patterns as well as facts like language change and the development of harmony systems.

References

Arsenault, Paul. 2012. *Retroflex Consonant Harmony in South Asia*. PhD dissertation, University of Toronto.

Brown, Jason Camy. 2008. *Theoretical aspects of Gitksan phonology*. PhD dissertation, University of British Columbia.

Gallagher, Gillian. 2010. Perceptual distinctness and long-distance laryngeal restrictions. *Phonology* 27(03): 435-480.

Gamta, Tilahun. 1989. *Oromo-English Dictionary*. AAU Printing Press.

Hansson, Gunnar Ólafur. 2001. *Theoretical and typological issues in consonant harmony*. PhD dissertation, University of California, Berkeley.

Hansson, Gunnar Ólafur. 2004. Tone and voicing agreement in Yabem. WCCFL 23: Proceedings of the 23rd West Coast Conference on Formal Linguistics, eds. B. Schmeiser, V. Chand, A. Kelleher and A. Rodriguez, 318-331. Somerville, MA: Cascadilla Press.

Hansson, Gunnar Ólafur. 2010. *Consonant harmony: long-distance interaction in phonology*. Berkeley, CA: University of California Press.

Kang, Yoonjung. 2011. Loanword phonology. *The Blackwell companion to phonology 4* (Vol. 4), eds. Marc van Oostendorp, Colin J. Ewan, Elizabeth Hume and Keren Rice, 2258-2282. Blackwell Publishing Ltd..

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2014. *Ethnologue: Languages of the World, Seventeenth edition.* Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Lloret, Maria-Rosa. 1995. The representation of glottals in Oromo. *Phonology* 12(02): 257-280.

McMullin, Kevin, and Gunnar Ólafur Hansson. 2013. Locality in long-distance phonotactics: evidence for modular learning. Talk presented at the 44th Meeting of the North East Linguistic Society (NELS 44), University of Connecticut, October 18-20, 2013.

Owens, Jonathan. 1985. *A grammar of Harar Oromo (Northeastern Ethiopia): Including a text and a glossary*. Vol. 4. Buske Verlag.

Ozburn, Avery. 2014. *Laryngeal harmony in Oromo*. MA Forum Paper, University of Toronto.

Ozburn, Avery, and Alexei Kochetov. 2013. Lezgian laryngeal harmony alternations. Poster presented at Phonology 2013, UMass Amherst, November 8-10, 2013.

Rose, Sharon, and Lisa King. 2007. Speech error elicitation and co-occurrence restrictions in two Ethiopian Semitic languages. *Language and Speech* 50(4): 451-504.

Rose, Sharon, and Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80: 475-531.