

TARGET-LIKE DISTRIBUTION OF NORWEGIAN LEXICAL PITCH ACCENTS IN SPONTANEOUS SPEECH PRODUCED BY L2 SPEAKERS

Guri Bordal Steien^{1,2}, Pernille Hansen¹

¹MultiLing/ILN, University of Oslo, ²University of Chicago
guri.bordal@iln.uio.no, pernille.hansen@iln.uio.no

ABSTRACT

In this paper, we present a study of tonal patterns of Accentual Phrases (APs) in spontaneous speech produced by ten Second Language (L2) speakers of Norwegian. The speakers are multilingual, and their dominant language is either Lingala or Swahili. Analyses of 1000 APs (100 per speaker) show that all the speakers produce the expected lexical pitch accent in a significant majority of the cases. Our initial hypothesis that the Lingala speakers, by virtue of speaking a language with lexical tones, have an advantage over the Swahili speakers in the acquisition of the Norwegian lexical pitch accents is not supported by the data.

Keywords: Lexical pitch accents, Second language acquisition, Norwegian, Lingala, Swahili

1. INTRODUCTION

Prosodic features of an L2, defined here as a language that is acquired after one or more other language(s), are claimed to be difficult to master [21, 22]. It has also been argued that prosody is particularly persistent in cross-linguistic influences [18], although it is not clear to what extent similarities and differences between a speaker's other languages and the target L2 play a role in the acquisition of L2 prosody [10, 20]. In this study, we examine tonal patterns of Norwegian APs produced by ten speakers with either Lingala or Swahili as dominant languages. The aim of the study is twofold. First, we seek to find out to what extent the speakers produce the expected lexical pitch accents of APs in spontaneous speech. Second, we are interested in whether phonological similarities between the speakers' dominant language and the target language play a role in the acquisition of prosody. Our hypothesis is that the Lingala speakers, by virtue of speaking a lexical tone language, have an advantage in the acquisition of the Norwegian lexical pitch accents over the Swahili speakers, whose dominant language does not have pitch contrasts at the word level. The data we have examined in this study, however, do not support the hypothesis: Our main finding is that all the speakers, regardless of their linguistic background, realize the

expected lexical pitch accent in a significant majority of the analyzed APs.

2. THE NORWEGIAN ACCENTUAL PHRASE

The AP is a prosodic constituent equal to or slightly larger than the word (see [11]). The tonal patterns of Norwegian APs are determined by the lexical pitch accent of the words they include, and tonal association interacts with word stress.

There is important variation in Norwegian prosody, but since the speakers of this study live in the South Eastern part of the country, we focus on East Norwegian (EN) here [8].

2.1 Word stress

As regards stress placement, EN polysyllabic words belong to one of the three following categories, of which the first is by far the most frequent [12, 13]: (i) Words that systematically have stress on the first syllable (*skole*, 'school'), (ii) words where stress is specified on some other syllable (*beholde*, 'keep'), and (iii) words that can receive stress either on the first syllable or on another syllable (*historie/historie* 'history'). There is important intra- and interspeaker variation in all of Eastern Norway with respect to the realization of the words in category (iii). Thus, the speakers in this study are likely to be exposed different realizations of them, a fact we have taken into account in our analyses (see Section 5.3).

The stressed syllable constitutes the left boundary of the AP, which includes all the following unstressed syllables. Unstressed syllables at the beginning of a word are included in the AP generated by the preceding word, and AP boundaries are not isomorphic to word boundaries if stress falls on another syllable than the first (*en lang historie* > [*en lang his*]_{AP1} [*torie*]_{AP2} 'a long history') [12, 13].

2.2 Lexical pitch accents

Norwegian has two lexical pitch accents, referred to as *accent 1* (/L*H/), and *accent 2* (/H*LH/) [14]. Every non-clitic word is linked to one of these accents. The first tones, /L*/ or /H*/, are associated with the stressed syllable, and the final /H/ tone is a boundary tone which is aligned to the last syllable of the AP. Monosyllabic words always have accent 1,

whereas the accent of polysyllabic words is generally not predictable, although there are some exceptions. Words with stress on the last syllable, for instance, are always associated with accent 1. Accent 2 is the most frequent, but the lexical specification of the accents is an object of debate (see for instance [14, 16]). Norwegian has some minimal pairs that are distinguished by their tonal pattern, such as *være* [²væ:rə](the verb ‘to be’) vs. *været* [¹væ:rə] (‘the weather’).

3. SPEAKERS

Five women and five men aged between 32 and 52 years (see Figure 3) participated in this study. They came to Norway as refugees from the Democratic Republic of Congo (DRC) in the beginning of the 2000s. When the recordings were made, they had lived in the country for around ten years. They all have one year of formal training in Norwegian, and use Norwegian on a regularly basis in their everyday life. Thus, their communicative competence in the language is good.

Common to all the speakers is that they spoke more than one other language before they learned Norwegian. Five of them come from the Kinshasa region where Lingala is the most spoken language, whereas the five others are from eastern parts of the DRC where Swahili dominates. Most of them also have some knowledge (at least passive) in other Congolese languages, such as Tshiluba, Kikongo or Mashi, although they always used them more rarely than Swahili or Lingala. Some speakers have also lived in other African countries for a period, and learnt the language spoken there. In addition, they speak French, the language of teaching in the DRC, and have studied English as a subject at school. Given that they are all multilingual, it is not straightforward to identify the role of some particular language in the acquisition of Norwegian; there is not one only *first language* that could have had an effect on the acquisition. When asked which language they considered as their *dominant language* (*i.e.* the language they speak better, the most often etc.), however, the speakers answered Lingala or Swahili. We therefore focus on the potential role of these languages in the acquisition of Norwegian here, although we cannot exclude influences from other languages (see Section 7).

4. LINGALA AND SWAHILI

Lingala has lexically specified tones, and maximal tonal density, *i.e.* every syllable is associated with a tone. Consequently, Lingala uses pitch to mark lexical contrasts, like Norwegian. Unlike

Norwegian, which has only two word melodies (accent 1 and accent 2), all kinds of word melodies are possible in Lingala [6, 9, 15].

Swahili has fixed word stress on the penultimate syllable of polysyllabic words. Stress is realized with different acoustic cues, but rising pitch is its main correlate [2]. Pitch is also used to mark focus and boundaries of prosodic constituents, but postlexical prosody does not affect the realization of stress [2, 3, 17, 23].

A main difference between Lingala and Swahili is that the former has lexical tones, while the latter has stress. We hypothesize that this difference might have had effects on the speakers’ acquisition of the Norwegian lexical pitch accents. If phonological similarities with the speakers’ dominant language facilitate the acquisition of L2 prosody, the speakers of Lingala are likely to have an advantage over the Swahili speakers in perceiving and producing melodic distinctions at the word level. A study comparing the perception of Norwegian lexical pitch accents by L1 speakers of a lexical tone language (Chinese) with speakers of a language without lexical tones (German) supports such a hypothesis: the Chinese speakers performed significantly better than the Germans [25].

5. CORPUS AND METHODS

The present study is based on analyses of a corpus with recordings of five minutes of spontaneous conversations for each speaker (one hour in total).

5.1. Context of conversations

Each speaker was recorded when he/she spoke alone with an L1 speaker of Norwegian, who did not share any of their other languages. The topics of the conversations were not predefined, but generally the speakers told about their experiences of coming to Norway as refugees. The style can globally be characterized as narrative. The recordings took place in the speakers’ homes.

5.2 Transcription, segmentation and correction

The recordings were transcribed orthographically in Praat [4], and segmented in words, syllables and phonemes by the Praat plug-in EasyAlign [7]. Since EasyAlign has not been trained for Norwegian, but for French, the automatic detection required extensive manual corrections. Syllables and phonemes were manually transcribed in Norwegian SAMPA. Errors in the pitch contour, such as octave jumps and intervening noises, were corrected with the software Anamor [1]. The software makes it possible to erase noises that appear in the

spectrograms, as well as to correct octave jumps by moving marked sequences of a pitch contour one octave up or down.

5.3 Annotation

We developed annotations in order to capture two aspects of the speakers’ productions that are relevant for the current study: (1) The tonal pattern of the realized APs, and (2) whether the realized tonal pattern corresponds to the expected pattern in EN.

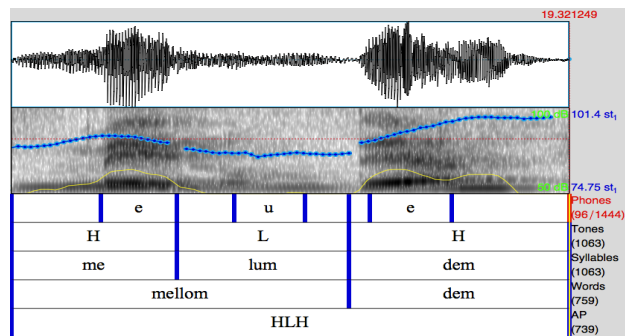
We selected tokens in the following way: Every non-clitic word was considered to potentially generate an AP. 100 utterance-internal APs of the transcribed conversation were analyzed for each speaker, a number we hypothesized would be sufficient to capture tendencies in the realizations of each speaker. We excluded words in contexts of hesitations, of overlapping turns or of poor pitch quality.

In order to identify the realized tonal patterns, we annotated the pitch of every syllable of the selected words, as well as of all the following syllables until the next potentially stressed syllable, according to the rules of AP formation in EN (see Section 2.1). We used the following labels: L (low), H (high), LH (rising), HL (falling). The annotation was carried out in two steps: First, we annotated the data solely based on visual interpretations of the spectrograms. Second, we corrected the first annotations on the basis of pitch measurements that were automatically made by Anamor. The software generates Excel documents with the mean pitch value and the value of the pitch rise/fall of each annotated syllable. To ensure that our annotations were consistent, we required that they respect the following four criteria: (1) Syllables on which pitch variations do not exceed one semitone were considered as having static pitch – thus as candidates for the labels L or H, (2) adjacent syllables must have a difference of at least one semitone to be labelled differently, e.g. a syllable annotated H has a pitch value that is at least one semitone higher than a preceding or following syllable annotated L, (3) syllables on which pitch variations exceed one semitone were labelled HL or LH, and (4) syllables with a pitch value that differed more than one semitone from both the syllable on its left and its right side received the same label as the syllable to which its pitch value was the closest. The choice of *one semitone* as the minimal difference between syllables annotated L and H is based on experiences from former studies, which have shown that this threshold makes it possible to single out regularities in languages where pitch is used for lexical contrasts (for a discussion, see [5, 6, 24]).

Finally, we created a new tier in Praat for the annotations of the tonal patterns of the APs. The labels we used here reflect the melodic pattern of the AP as a whole (for instance if it contains four syllables with the labels HL, L, L, H, the AP was annotated HLH). If the speakers produced the expected patterns, the following labels were used: HLH (AP correctly realized with accent 2) and LH (AP correctly realized with accent 1). The APs that had other patterns than the expected ones received one of the following labels: HLH>LH (AP containing an accent 2 word realized with the pattern of accent 1), LH>HLH (AP containing an accent 1 word realized with the pattern of accent 2), LH>X (AP containing an accent 1 word realized with another tonal pattern) and HLH>X (AP containing an accent 2 word realized with another tonal pattern). By the *expected pattern*, we generally mean that of the lexical pitch accent associated to the word according to the dictionary of EN pronunciation [26]. There is, however, some variation in EN pronunciation that is not captured by the dictionary, and we used our intuition as first language speakers to check whether another pattern could also be possible in cases of doubt.

We counted the labels with the function “count labels” in Praat.

Figure 1: Example of an annotated AP in Praat, which is correctly realized with accent 2 (*mellom dem*, ‘between them’)



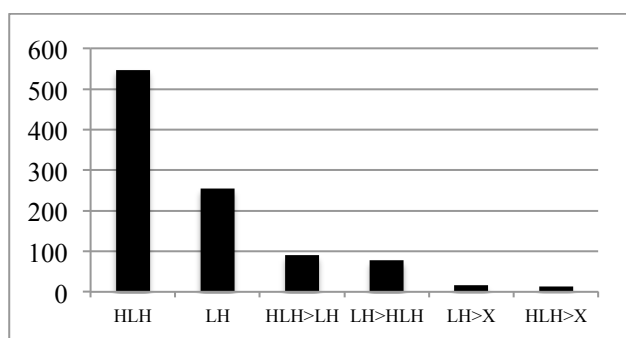
6. RESULTS

The distribution of tonal patterns in the corpus is given in figure 2. As 802 of the 1000 analyzed tokens were realized with the expected tonal pattern, the speakers are statistically more likely to realize an AP with the expected lexical pitch accent than with another tonal pattern ($\chi^2(2)=372.78$, $p<.005$). This is also true for each speaker analyzed individually (Fisher’s exact tests: $p<.005$ for every speaker). The distribution of melodies for each speaker is shown in figure 3.

As regards the APs which are not realized with an expected pattern, it is interesting to note that they are

significantly more often realized with the pattern of the other pitch accent (HLH>LH or LH>HLH, in sum accounting for 85 % of the unexpected patterns) than with patterns that do not exist in the Norwegian tonal system (HLH>X and LH>X) ($\chi^2(1)=97.58$, $p<.005$). Finally, there is a bias towards accent 2 (HLH), the most frequent pattern in Norwegian: LH APs are less frequent in the data, and they are more often realized with the other pitch accent than HLH APs ($\chi^2(1)=10.75$, $p<.005$); whereas 22 % of the LH APs (78 of 349) are realized as HLH, only 14 % of the HLH APs (91 of 651) are realized as LH.

Figure 2: Distribution of tonal patterns in the corpus. The y-axis indicates the number of APs.

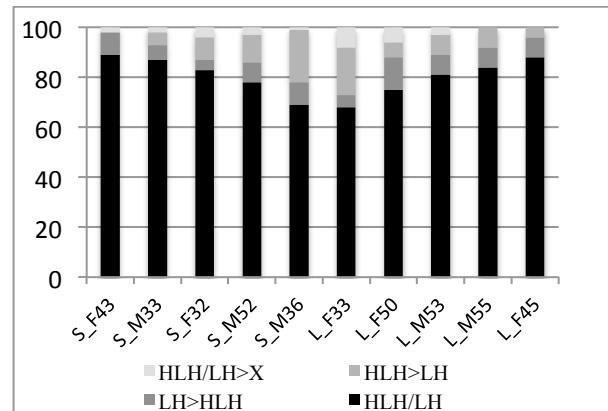


6.1 Lingala speakers vs. Swahili speakers

The data do not support the hypothesis that Lingala speakers have an advantage over Swahili speakers in acquiring the Norwegian lexical pitch accents. In fact, there is a small tendency in the opposite direction: the Lingala speakers produce 108 APs with unexpected patterns, whereas the Swahili speakers produce 94. This difference, however, is not statistically significant ($\chi^2(1)=0.51$, $p=.48$). There are no significant differences in the types of unexpected patterns between the groups either ($\chi^2=1.04$, $p=0.8$): X patterns are rare within both groups, and they have the same bias towards HLH.

There were, however, significant differences between the individual speakers as regards the number of unexpected patterns ($\chi^2(9)=35.83$, $p<.005$). These differences are due to the realizations of three speakers: two speakers produced more unexpected patterns than the others (S_M36 and L_F33), and one produced less unexpected patterns (S_F43) (see Figure 3). If we exclude these speakers from the analysis, the differences between the speakers are no longer significant ($\chi^2(7)=12.18$, $p=.09$).

Figure 3: Distribution of patterns in each speaker. The codes indicate the language (L=Lingala, S=Swahili), the gender, (F/M) and the age of the speaker.



7. DISCUSSION AND CONCLUSION

In this paper, we have presented a study of tonal patterns of Norwegian APs produced by ten L2 speakers. We have shown that all the speakers realize the expected lexical pitch accent in a significant majority of the analyzed APs, and that there are no significant differences between the speakers of Lingala and Swahili. These findings are interesting for at least two reasons: First, it has been pointed out in the literature that speakers often have difficulties acquiring the prosody of an L2, even after several years of exposure [22]. The results presented here, however, show that speakers have reached a high level of proficiency in the Norwegian tonal system, although there is a slight tendency of overgeneralization of accent 2, the most frequent of the EN lexical pitch accents. Other aspects of their productions, such as scaling, alignment, correlates of stress, pragmatic use of prosodic cues, etc., should nevertheless, be studied to get a full overview of the speakers' knowledge of Norwegian prosody. Second, contrary to our hypothesis, we found no differences between the Lingala and the Swahili speakers. In this respect, our study presents a clear limitation; we only have five speakers per group, and the results could have been different in a bigger sample. On the other hand, the fact that all the speakers already spoke languages with different prosodic systems before they acquired Norwegian could possibly have played a more important role in the acquisition process than similarities and differences between target and dominant language.

7. REFERENCES

- [1] Avanzi, M., Lacheret-Dujour, A., Victorri, B. ANALOR. A Tool for semi-automatic annotation of French Prosodic Structure. *Proceedings of Speech Prosody, Campinas (Brazil)*, 119-122.
- [2] Ashton, E. O. 1947. *Swahili Grammar (Including Intonation)*. London: Longmans.
- [3] Burt, F. 2011. *Swahili Grammar and Vocabulary*. München: LINCOM
- [4] Boersma, P., Weenink, D. 2011. Praat, version 5.3.51 <http://www.praat.org>
- [5] Bordal, G. 2012. *Prosodie et contact de langues: le cas du système tonal du français centrafricain*. Ph.D. thesis, Universitsty of Oslo/Université Paris Ouest Nanterre la Défense.
- [6] Clements, G. N. 1979. The description of terraced-level tone languages. *Language* 55, 536–558.
- [7] Goldman, J.P. 2008. EasyAlign: A semi-automatic phonetic alignment tool under Praat. <http://latlcui.unige.ch/phonetique/>
- [8] Fretheim, T, van Dommelen, V.A. 2012. A pragmatic perspective on the phonological values of utterance-final boundary tones in East Norwegian intonation. *The Linguistic Review* 29, 663–677.
- [9] Guthrie, M. 1940. Tone ranges in a two-tone language (Lingala). *Bulletin of the School of Oriental Studies, University of London* 10, 469–478.
- [10] Jilka, M. 2007. Different manifestations and perceptions of foreign accent in intonation. In: Trouvain, J., Gut, U. (eds), *Non-Native Prosody. Phonetic Description and Teaching Practice*. Berlin: Mouton de Gruyter. 145–167.
- [11] Jun, S.-A. (ed) 2014. *Prosodic Typology II: The Phonology of Intonation and Phrasing*. New York: Oxford University Press.
- [12] Kristoffersen, G. 2000. *The Phonology of Norwegian*. Oxford: Oxford University Press.
- [13] Kristoffersen, G. 2002. *Norsk prosodi*. Unpublished manuscript. Nordisk institutt, Universitetet i Bergen.
- [14] Kristoffersen, G. 2006. Markedness in Urban East Norwegian tonal accent. *Nordic Journal of Linguistics* 29, 95–135.
- [15] Kukanda, V. 1983. *L'emprunt français en lingála de Kinshasa: quelques aspects de son intégration phonétique, morphologique, sémantique et lexicale*. Tübingen: G. Narr.
- [16] Lahiri, A., Wetterlin, A., Jönsson-Steiner, E.. 2006. Scandinavian lexical tone: Prefixes and compounds. In: Bruce, G., Horne, M (eds), *Nordic Prosody: Proceedings of the IXth Conference*. Frankfurt am Main: Peter Lang, 167–173.
- [17] Lodhi, A. 2004. Strategies of emphasis intensity in Swahili. *Africa & Asia* 4, 142–150.
- [18] Mackey, W. F. 2000. The description of bilingualism. In Wei, L. (ed) *The Bilingualism Reader*. Oxford: Routledge, 26–54
- [19] Maw, J., Kelly J. 1975. *Intonation in Swahili*. London: SOAS.
- [20] Mennen, I. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32, 543–563.
- [21] Mennen, I. 2007. Phonological and phonetic influences in non-native intonation. In: Trouvain, J., Gut, U. (eds), *Nonnative Prosody: Phonetic Descriptions and Teaching Practice*. Berlin: Mouton De Gruyter, 53–76
- [22] Mennen, I. 2015. Beyond segments: towards a L2 intonation learning theory in prosody and language. In: Delais-Roussarie, E., Avanzi, M., Hermet, S. *Contact L2 Acquisition, Attrition and Languages in Multilingual Situations*. Berlin Heidelberg: Springer, 169-186
- [23] Mpiranya, F. 1995. *Swahili Phonology Reconsidered in a Diachronical Perspective: the Impact of Stress on Morphonemics and Syllable Structure*. Köln: Rudiger Köppe.
- [24] Steien, G. B. 2014 Andrespråktoner gjennom første-språkprosodi: en studie av ordmelodier i norsk spontan tale produsert av personer med lingala og swahili som S1. *Nordand. Nordisk tidsskrift for andrespråksforskning* 9(2), 9–30.
- [25] Van Dommelen, W. A., Husby, O. 2009. Perception of Norwegian word tones by Chinese and German listeners. In: Watkins, M., Andreia A., Rauber S., Baptista, B. O. (eds), *Recent Research in Second Language Phonetics/Phonology: Perception and Production*. Cambridge: Cambridge Scholars Publishing, 318–331.
- [26] Vanvik, A. 1985. Norsk uttaleordbok. Fonetisk institutt, Universitetet i Oslo. <https://www.scribd.com/doc/122451989/Norsk-Uttaleordbok-av-Arne-Vanvik>

8. ACKNOWLEDGEMENTS

This work was partly supported by the Research Council of Norway through its Centres of Excellence funding scheme, project number 223265.