LANGUAGE AT THE UNIVERSITY OF ESSEX (LangUE) PROCEEDINGS 2018

TABLE OF CONTENTS

INVESTIGATING ALGERIANS' CODESWITCHING ON FACEBOOK PAGES

FIRDOUS ABDELHAMID
(Reading University)

Abstract

Codeswitching is a common linguistic practice in multilingual speech communities. Motivated by the literature on the Algerian multilingual speech community that affirms the existence and the habitual codeswitching behaviour of its speakers in spoken discourse (Bagui 2014), this study investigated the extent to which Algerians codeswitched in other genres of interactive discourses, mainly those mediated through computers. Two Algerian Facebook pages were selected and data was collected from a total of 120 Facebook posts and 16355 comments. The data analysis showed that one third of the collected data were messages that contained instances of codeswitching while two thirds of the collected data were monolingual. The study found that codeswitching between Algerian-Arabic and French and between Algerian-Arabic and Modern Standard Arabic were the most frequent cases of codeswitching while the use of English was limited.

1. Introduction

Having more than one language or language variety in one's linguistic repertoire will inevitably lead to switching between them when speaking. Code will be used as a neutral term to refer to the languages and language varieties of interest. As such, switching between codes or codeswitching (henceforth CS) is a common linguistic behaviour among multilingual speech communities. This study examined whether Algerians that were brought up in a multilingual speech community codeswitched when communicating online, namely when using the social media platform, Facebook. The present paper is structured into four main sections. Sections 2 and 3 represent a brief review of the related literature that has been investigated on CS on Facebook and the multilingual situation of the Algerian speech community. Sections 4 and 5 explain the data collection procedures and the data analysis process respectively. Finally, section 6 presents the results and discussion of the study in an attempt to answer the research questions.

2. Codeswitching on Facebook

It has long been acknowledged that spoken discourse is a unique source of spontaneous communication that contains instances of CS. Therefore, most research on CS is based on analysing spoken discourse. Attention, however has recently shifted to other genres of discourse, namely Computer-Mediated Discourse (CMD) (Androutsopoulos 2013) as it has been observed to be a platform where CS is pervasive (Friedrich & de Figueiredo 2016). In addition to CS, CMD shares other characteristics with spoken discourse that include the typographic practices that resemble prosody and intonation and turn-taking (Herring 2010). According to Barasa (2015), CS in CMD is not spontaneous but rather a conscious and planned practice. Conscious CS data is suitable for research that is concerned with identity issues, style expressions (Dorleijn & Nortier 2009) and discursive functions. This study is part of a larger project that is interested in explaining how CS on Facebook is used to achieve such ends. It is a preliminary study that explores whether Facebook could be a source of quality data and provide sufficient instances of CS to conduct the larger project. This is

because Facebook could be an asynchronous mode of CMD. Asynchronous modes are notorious for being a platform where CS is disfavoured (Paolillo 2011). The assumption is that when people have more time to produce their messages, they will revise them to remove any instances of CS in order to achieve what is deemed as 'correct' language since CS has always been synonymous with the corrupt use of language (Sebba 2012). Previous research on CS on Facebook found that CS is a natural linguistic behaviour that is used by bi-multilingual speakers as a strategy to structure their interactions (cf. Auer 1984), to achieve communicative ends (Halim & Maros 2014) and to project different identities (Sophocleous & Themistocleous 2014). CS on Facebook serves pragmatic and discourse functions that include quoting others and showing speakers emotions. It is used as a strategy when no equivalent of a given word is found in the language that is used to write the message or if it is too long and time consuming to type (Dąbrowska 2013, Halim & Maros 2014). In addition, CS is used for addressing different audiences (Eldin 2014) and to index discursive and social identities that include serious or funny, regional or global and youth (Sophocleous & Themistocleous 2014).

3. Linguistic Context of Study

In Algeria, Algerian-Arabic (AA) is a non-standard regional variety that is spoken by most of Algerians (Chemami 2011). It exists alongside the standard variety referred to as Modern-Standard Arabic (MSA). Algerians use AA for spoken discourse and MSA for written discourse, education and formal occasions. The educational system in Algeria promotes the teaching of two foreign languages: French and English. Baghi (2014: 86) describes the Algerian linguistic situation as 'very intricate'. The intricacy lies in the co-presence of these codes in one speech community which basically promotes the prevalence of CS. She maintains that because of historical reasons, namely French colonialism, French is not only present but deeply rooted in the Algerian society to an extent that upon exposure to natural and spontaneous speech of Algerians, one notices the habitual switching back and forth between AA and French. These are not the only existing pairs of switching as, according to her, the Arabisation process, of replacing French with Arabic in vital domains of the country after independence, produced a group of educated Arabophones who switch between AA and MSA. Baya Essayahi and Kerras (2016) argue that this Algerian speech which is a mixture of AA and French and/ or MSA terms is innate and should not be treated as intentional CS but rather as a distinct variety called Algerian. Algerian, according to Baya Essayahi and Kerras (2016), is a developing language that needs to be codified and even taught in schools. A proposal that the Algerian ministry of education has rejected.

Published studies on CS within the Algerian context are limited. Benguedda (2015) who was interested in CS to French conducted a study to explore the social factors that promote CS. Following the researcher's observations that almost all Algerians used French items with Arabic ones regularly, she raised the question of which social factors are involved in the practice of CS. She collected tape-recordings of 10 families from two districts of different status, i.e. high versus low. The analysis revealed that participants coming from rural areas tended to speak mainly in AA when interviewed whereas those coming from urban areas codeswitched between AA and French even if they were uneducated. It was found that socio-economic status seemed to be a factor for CS but not gender. Both male and female participants codeswitched. However, CS is not exclusive to spoken discourse and can be observed on advertising placards all around Algeria. Boukreris and Bouchegra (2016) noted that advertising companies use CS as a strategy to promote their products. These companies try to narrow the gap between the consumers and themselves by using language

that is familiar in a consumer's daily life. Inserting AA terms in a sentence written in French raises emotions of loyalty towards local products which drives consumers to buying them.

The present study takes a different direction and raises the question of whether this habitual CS practice is observed online. It is interested in the frequency of conversational written CS on Facebook. In previous studies, very few have tackled the frequency aspect of CS in Algeria which is possibly due to the difficulty of distinguishing between CS and other similar processes like borrowing. In this study, borrowing is defined as the act of assimilating words (Poplack 1980) of different origins into AA to an extent of being intelligible to AA monolinguals. Any use or insertion of words of other origins that necessitates bi-multilingual knowledge is approached as intentional CS which is worth the analysis because of the social significance (cf. Meyers-Scotton 1988) it carries and not as innate CS as argued by Baya Essayahi and Kerras (2016). Accordingly, the research questions of this study are:

1. How frequent is Algerian CS on Facebook pages?

2. What codes are involved in CS instances?

4. Data collection

Data for the present study was collected from two Facebook pages that are popular among the Algerian online community. Facebook was selected as a source of data because it is one of the most popular social media platforms used by Algerians. When asked about their preferred social media site, students chose Facebook as oppose to five other sites which include YouTube and Twitter (Boumarafi 2015). Also, Algerians are the fourth biggest users of Facebook in Africa (Statistics from 2016). The selected Facebook pages are referred to as Algerian Facebook Page 1 (AFP1) and Algerian Facebook Page 2 (AFP2). They were selected based on three criteria. First, they have a public status so any user of Facebook can access their publications and write comments on them. Second, they are popular among Algerians as AFP1 is liked by almost 1.5 million users and AFP2 is liked by almost 1.4 million users. Finally, both are platforms where members can use whichever code they wish to. There are no language use regulations and based on the researcher's observations more than one code was being used. In reading the 'About' section that the administrators provided in each page, it is clear that AFP1 is focused more on entertainment purposes whereas AFP2 is more informative especially in relation to reporting the daily issues of Algerians as well as discussing political and educational matters.

To collect the data, the news feed of both pages were observed which resulted in the selection of 60 posts from each page. Only posts where language was used were collected; posts that shared links, photos or videos and those that revealed the identity of the page were disregarded. To extract the chain of comments underneath each post, the researcher relied on the software 'Facepager' and the online tool 'the crazy comment export', both of which extracted the comments from Facebook into tables. A total of 16355 comments were extracted into tables but only 11920 were analysed and used in the study. This is because the remaining 4569 were comments which included emojis, emoticons and/or images and did not make use of language.

5. Data Analysis

To answer the first research question, each line of the messages in the tables was read repeatedly to decide on the code(s) that it was written in and whether it displayed any CS. This was not an easy task to accomplish as there was no explicit set of rules to distinguish between some ambiguous cases and some items were difficult to categorise as belonging to AA or French/MSA. Previous researchers who encountered this problem dealt

with it in relation to the level of assimilation of words. For Ahmed-Sid (2009), Bengedda (2016) and Slimane (2014), any word that has a foreign origin but is morphologically, syntactically and phonologically assimilated into AA, was treated as belonging to AA and not considered as CS. Unlike CS, borrowed items become part of the recipient code, used by the whole speech community and are transmitted to the following generations. Examples of this include *taksifon* 'public call box' and *klinik* 'the clinic' (Bengedda 2016). This view follows the ideas of Poplack and her associates that distinguished between CS and borrowing based on assimilation levels (Poplack 1980). Yet, there exists a different position in the literature which relies on frequency as opposed to assimilation as the distinguishing characteristic (Myers-Scotton 1993). Taking all this into consideration, a list of encountered borrowed items in the data was accumulated based on the researcher's knowledge as a native speaker of AA that such items are assimilated into AA and are frequently used by the entire speech community. As such, in the analysis, these items are treated as AA words and not cases of CS.

Another problem that arose due to the medium of communication relates to how words from AA and/or MSA can be undistinguishable when written down using the Arabic script. For example, the phrase 'I wrote' is written in AA as كتبت which is the same in MSA كتبت. The only difference between them is the pronunciation whereby the phrase is pronounced *ktebt* in AA but *katabtou* in MSA. To overcome this problem and to ensure the credibility of the frequency rates, when such words were found in an AA message they were considered as AA items and not case of switching to MSA and vice versa.

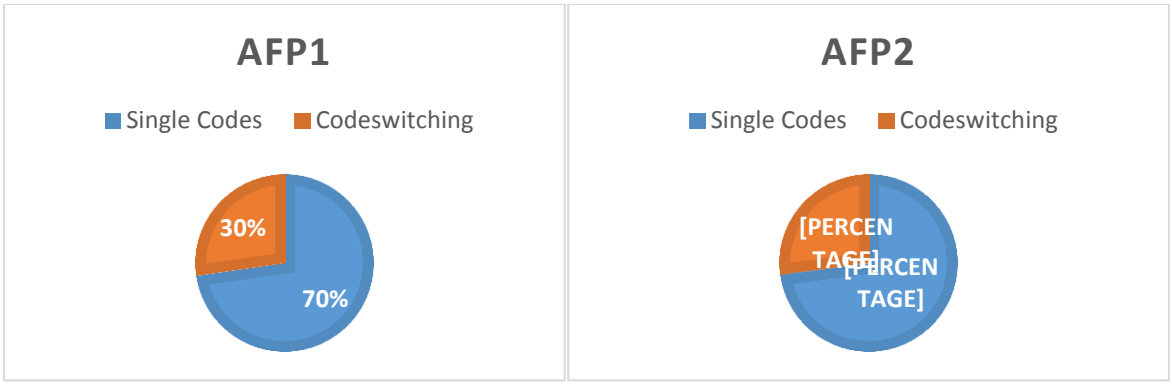
AA is represented in the data in normal font, French in **bold**, English in *italics* and MSA <u>underlined</u>. The same applies for the English translation underneath the phrase so that the reader can follow the examples clearly.

6. Results and Discussion

6.1. Frequency of codeswitching

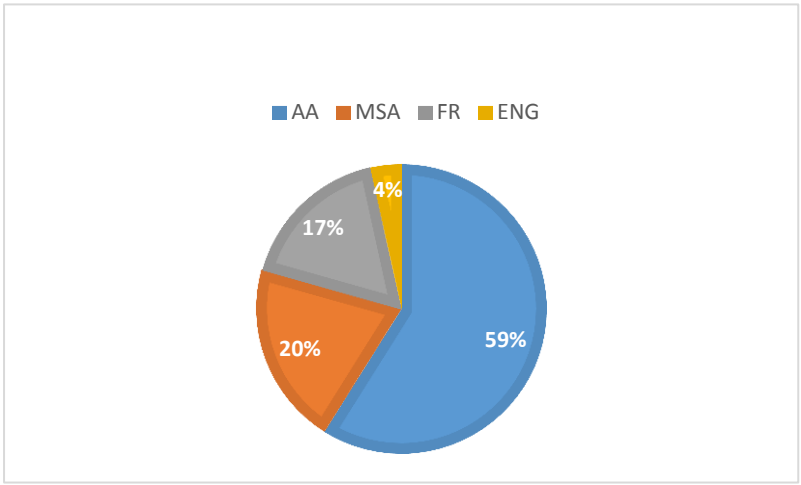
Quantitative analysis of the messages revealed that a third of the collected data were messages that contained CS instances and the rest were monolingual, Figure 1. This is the case for both Facebook pages, AFP1 and AFP2. Writing comments on Facebook posts could mean writing very short and fragmented messages and the fact that they were written in more than one code rises interests to study them. It seems that CS is triggered in such kinds of communications through Facebook where it can be an asynchronous mode of communication that was predicted to disfavour CS (Paolillo 2011). The presence of CS in these two Facebook pages reflect the multilingual speech community of Algeria and confirms that CS is a natural phenomenon in language contact situations which can be observed online. This observation is shared with previous studies investigating CS on Facebook (e.g. Dabrowska 2013, Eldin 2014, Halim & Maros 2014, Sophocleous & Themistocleous 2014).

Figure 1: Frequency of codeswitching



The results in Figure 1 show that the frequency of CS in AFP1 and AFP2 is very similar with 30% in AFP1 and 27% in AFP2. The difference is the codes that users codeswitch between. For the messages that were written in single codes, the analysis revealed that all codes available to Algerians are used on Facebook but with varying frequency rates. These codes are Algerian-Arabic (AA), Modern Standard Arabic (MSA), French (FR) and English (ENG).

Figure 2: Frequencies of messages written in single codes



The expected finding to emerge from the data was the wide use of AA. AA was the most widely used code to write messages in both Facebook pages with a frequency of 59% out of the total number of messages written in single codes. This is because AA is the mother tongue of the majority of Algerians and as Barasa (2015) explains, the mother tongue of the majority is used for writing comments to ensure the maximisation of one's audience since Facebook users are not fully aware of who is reading their comments. MSA is the second chosen language for writing comments followed by both French and English with frequencies of 20%, 17% and 4% respectively. This echoes the code position in Algeria with MSA as the official language and French and English as the first and second foreign languages respectively. The findings of this study on the use of French on Facebook reinforces the argument that despite all the Arabisation efforts and policies to replace the French language in Algeria by either Arabic or English because of its colonial connotations (Bensafi 2002), French is still powerfully present in the oral communication of Algerians (Mostari 2011) as well as in its written form. Some of the negative attitudes towards the use of French relate to

issues of colonisation and a loss sense of identity (Ahmed-Sid 2008). The following comment was posted by a member of AFP2 to a post of an article in French, example (1). He expresses how it irritates him to see anything written in French asking 'when are we going to become Algerians' which demonstrates how he relates the use of French to Algerian identity. For him, being Algerian means using AA or at least not using French.

(1) على بالكم اني كي نشوف حاجة مكتوبة بالفرونسي تحكمني الغمة ... وينتا يرجع هذا الشعب جزايري

'Do you know that I really get pissed off when I see something written in French… when are we going to become Algerians.'

The use of English in writing comments is clearly not the first choice for the majority as only 2.63% of messages were written entirely in English. Nonetheless, it is still noteworthy to observe its use on Facebook pages although statistics show that only 7% of Algerians are capable of communicating in English (Benrabah 214). Indeed, in a study by Chemami (2011), 74.3% of participants stated that they used English rarely or very rarely in their daily communication. However, they were still interested in improving their English. One member of AFP1 shared this positive attitude towards the use of English as he/she posted:

(2) *what if ' AFP1 ' is in english ! that will be so fucking awesome*

The findings become more interesting when code use in each page were compared separately. While MSA is the second most used code in AFP2, it is French that is the most widely used code in AFP1. One explanation for this could be the topics that these Facebook users correlate to the codes. As previously mentioned, AFP1 posts are entertaining posts and allow more casual discussions favouring the use of AA and French but in the case of AFP2, discussions are of a more formal nature. Most of the posts in AFP2 are about education, politics and religion, all of which are usually addressed in MSA. In other words, the language of instruction in classrooms from primary schools to high schools is MSA, and it is also the language of religion, politics and the media. Similar findings were found in a study by Sophocleous and Themistocleous (2014) which examined the use of the standard Greek variety on Facebook whereby standard Greek was used for more formal and serious communicative purposes. Although it was not the focus of this study, findings suggest that there is a correlation between the topics discussed and the codes used. This idea was refuted by Hinrich (2006) who argued against any generalisation that could be made between topics and language use as users are free in their linguistic choices.

6.2. Codes used in the codeswitching instances

With regards to the second research question, and adapting from Barasa's (2015) terminology, the study revealed that users used both bilingual and trilingual codeswitching in writing their comments. Bilingual codeswitching is the switching between two codes and trilingual codeswitching is the switching between three codes. Most of the comments were written in bilingual codeswitching as only 6% of comments contained trilingual codeswitching.

Cases of bilingual codeswitching included among others switching between the pairs AA and MSA, AA and FR and AA and ENG. The most frequent pair in the data was AA and FR with 46.05 %. Switching between AA and French was also most prevalent in SMS exchanges between young Algerians with a proportion of 41.5% (Mostari 2009). Examples from the data revealed that users of Facebook codeswitched between AA and French because it entails 'less effort' (Barasa 2015). When AA words are judged to be too long or time consuming to type, writers of comments tend to switch to French to type abbreviations of French words instead. In the selected comments in example (3), writers used the items *tjr* and *svp* that are abbreviations for *toujours* 'always' and *s'il vous puisque* 'please' instead of the equivalent words in AA that could be *dima* and *t3ich* respectively.

(3) a. 7na **tjr** hakdha
'we are **always** like this.'

b. wache kayan **svp**
'whats going on **please**.'

The findings indicated that switching between AA and MSA was more frequent in AFP2 than AFP1 and this could be due to the topics of the posts as previously mentioned. For instance, the comment in example (4) was written to a post about the high school final exam, the baccalaureate. The post provided some tips to help answer exam questions and the writer of the comment stated that it was necessary to include some terms in writing the paragraphs. However, because these terms were related to education there was a switch from AA to MSA.

(4) جغرافيا لازم تقول الموقع الاستراتيجي وخصوبة التربية ودوك الصوالح
'In geography, you have to use Strategic location and Soil fertility and terms like that.'

It is worth noting the presence of CS between AA and English since previous investigations of CS in Algeria reported limited use of English in spoken discourse and tended to disregard its analysis (e.g. Ahmed-Sid 2008). Similarly, Mostari (2009) found that the use of English in SMS exchanges was also very limited with only 0.2%, of the 50 SMS texts examined, contained English items.

In some cases of the collected data, users wrote comments in English and switched to AA for 'specific vocabulary' (Dąbrowska 2013) as shown in example (5). The writer of this comment, in answering a question on whether his or her dreams became true, wrote the answer in English and then switched to AA to include a religious phrase of gratitude. *Elhamdoulilah* 'praise and thanks are due to Allah' is of MSA origin and has been assimilated into AA and is widely used for thanking god no matter what code the speaker is using. This is possibly a way to ensure that the religious connotations of the expression would not be lost in translation.

(5) *Yes* Elhamdoulilah *...but I have a bigger one couldn't make it true yet hope soon* :D
'*Yes* praise and thanks are due to Allah *...but I have a bigger one couldn't make it true yet hope soon* :D.'

Cases of trilingual CS include the switching between AA, FR and ENG. Switching between these codes was the most frequent case of trilingual switching with a frequency of

4.16%. The example below is a comment written in this type of CS. In example (6), the writer answered the previous question in French stating that the dream became true but used English to elaborate that although the dream was true, there were other dreams that he or she was working on. AA items were also inserted mainly due to vocabulary specifications. Two religious expressions were used; the first was to thank god for realising the dream and the second was to hope that god would help the writer in achieving more.

(6) **Oui** hamdoullah **mais** *the best is still to come* incha'aallah
'**Yes** praise and thanks are due to Allah **but** *the best is still to come* if Allah wills.'

7. Conclusion

CS has enjoyed great attention when found in spoken discourse but is an under-researched area in CMD and written discourse. This study is an attempt to contribute to the literature of CS in writing. It has sought to explore the frequency of Algerian CS on Facebook which will enable a larger, more focused qualitative study to examine in detail these CS instances. Data was collected from two Facebook pages, AFP1 and AFP2. Results revealed that similar to Algerian spoken communication, CS is observed in Facebook communications as well. One third of the collected data were messages that were written in more than one code. Mostly, these were cases of bilingual CS as cases of trilingual CS were limited. Most users codeswitched between AA and French and between AA and MSA. These findings strongly indicate that CS is a natural phenomenon and is practised even on social media platforms.

References

Ahmed-Sid, H. (2008). Code-Variation among Algerian University Students: A Case Study. Unpublished Doctoral Thesis. Mentouri University, Constantine.

Ahmed-Sid, H. (2009). Code-switching and borrowing in Algeria. *Revue Sciences Humaines*, (32), 97-107.

Androutsopoulos, J. (2013). Code-switching in computer-mediated communication. In Herring, S., Stein, D. & Virtanen, T. (eds.), *Pragmatics of Computer-mediated Communication*, 667–694. Berlin: Mouton de Gruyter.

Auer, P. (1984). *Bilingual conversation*. John Benjamins Publishing.

Bagui, H. (2014). Aspects of diglossic code switching situations: A Sociolinguistic Interpretation. *European Journal of Research in Social Sciences*, 2(4), 86-92.

Barasa, S. (2015). Spoken Code-Switching in Written Form? Manifestation of Code-Switching in Computer Mediated Communication. *Journal of Language Contact*, 9(1), 49-70.

Baya Essayahi, M. L., & Kerras, N. (2016). A Sociolinguistic Study of the Algerian Language. *Arab World English Journal* (3). 141-154.

Benguedda, A. (2015). Consequences of language contact: Case of social factors that affect code switching. *Bulletin VALS-ASLA*, 3, 13-29.

Benrabah, M. (2014). Competition between four "world" languages in Algeria. *Journal of World Languages*, 1(1), 38-59.

Bensafi, Z. (2002). Bilingualism in Algeria and the making of the nation. *Second International Symposium on Bilingualism*.

Boukreris, L., & Bouchegra, A. (2016) The Language of Advertising: The Case of Code-Switching and Code-Mixing in the Algerian Advertising Scenery. 40-50.

Boumarafi, B. (2015). Social Media Use in Algerian Universities: University of Constantine 2 Case Study. *IAFOR Journal of Education*. 31-41.

Chemami, M. A. (2011). Discussing plurilingualism in Algeria: the status of French and English languages through the educational policy. *International Journal of Arts & Sciences*, 4(18), 227-234.

Dąbrowska, M. (2013). Functions of code-switching in polish and Hindi Facebook users' posts. *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, (130), 63-84.

Dorleijn, M., & Nortier, J. (2009). Code-switching on the Internet. In: Bullock, B. &. Toribio, A. J. (eds.), *The Cambridge handbook of codeswitching*, 127–141. Cambridge: Cambridge University Press.

Eldin, A. A. T. S. (2014). Socio linguistic study of code switching of the Arabic language speakers on social networking. *International journal of English linguistics*, 4(6), 78-86.

Friedrich, P., & de Figueiredo, E. H. D. (2016). *The Sociolinguistics of Digital Englishes*. Routledge.

Halim, N. S., & Maros, M. (2014). The functions of code-switching in facebook interactions. *Procedia-Social and Behavioral Sciences*, 118, 126-133.

Herring, S. C. (2010). Computer-mediated conversation: Introduction and overview. *Language@ internet*, 7(2), 1-12.

Hinrichs, L. (2006). *Codeswitching on the web: English and Jamaican Creole in e-mail communication* (Vol. 147). John Benjamins Publishing.

Mostari, H. A. (2009). What do mobiles speak in Algeria? Evidence from SMS language. *Current Issues in Language Planning*, 10(4), 377-386.

Mostari, H. A. (2011) new sociolinguistic insight into language contact phenomena: evidence from Arabic/French languages pair. *Language Discourse &Society*, 126-139.

Myers-Scotton, C. (1993). Common and uncommon ground: Social and structural factors in codeswitching. *Language in society*, 22(4), 475-503.

Myers-Scotton, C. (2002). Code-switching as indexical of social negotiations. In Li, W. (ed.), *The bilingualism reader*, 140-166. Routledge.

Paolillo, J. C. (2011). Conversational" codeswitching on Usenet and Internet relay chat. *Language@ Internet*, 8(3).

Poplack, S. (1980). Sometimes i'll start a sentence in spanish y termino en espanol: toward a typology of code-switching1. *Linguistics*, 18(7-8), 581-618.

Sebba, M., (2012). Researching and theorising Multilingual texts. In Sebba, M., Mahootian, S. & Jonsson, C. (eds.), *Language mixing and code-switching in writing: Approaches to mixed-language written discourse*, 1-26. Routledge.

Slimane, H. M. K. (2014). Bilingualism and biculturalism: Exploring youth language, culture and identity in today's Algeria. *International Journal of Research in Humanities, Arts and Literature*, 2(2), 11-18.

Sophocleous, A., & Themistocleous, C. (2014). Projecting Social and Discursive Identities Through Code-Switching on Facebook: The Case of Greek Cypriots. *Language@ Internet*, 5(11).

Firdous Abdelhamid
Department of English Language and Applied Linguistics
University of Reading
UK

f.abdelhamid@pgr.reading.ac.uk

LIVONIAN – A CASE FOR A HYBRID MODEL OF MORPHOLOGICAL DESCRIPTION?

ZEPRINA-JAZ AINSWORTH
(University of Oxford)

Abstract

Blevins (2006) identifies and describes a dichotomy of morphological models, referring to the two approaches as 'constructive' and 'abstractive'. Earlier work has demonstrated convincingly that languages with fusional structure cannot be accommodated by constructive models which posit the morpheme as the smallest meaningful unit of phonological form. However, despite the unit-agnostic nature of modern abstractive approaches, little importance is afforded to sub-word units which demonstrate an isomorphic function~form mapping. In this paper, I suggest that languages which exhibit both fusional and agglutinative structure are better accommodated by a model which combines essential features of both abstractive and constructive approaches to grammatical description.

1. Introduction

This paper aims to argue for a hybrid model of morphological description, combining aspects of the 'abstractive' and 'constructive' models of grammatical description.[1] Such a model is essential for the accommodation of the inflexional patterns found in the Livonian morphological case system. Whilst the Item and Pattern[2] perspective captures the implicational patterns that hold across the Livonian inflexional case system, the stability of sub-word mapping of function and form in certain instances is better captured by a constructive description.

The structure of this paper is as follows: §2 will introduce the main tenets of the abstractive and constructive approaches to morphological description, highlighting the differences between the two models which make it necessary to recognize these two as dichotomous; in §3 the Livonian data will be presented; earlier (constructive) work on the description of Livonian inflexion classes will then be briefly outlined in §4 and an alternative, abstractive account will be offered in §5, whilst §6 will be concerned with presenting the 'new' hybrid constructive-abstractive model in reference to the Livonian inflexional case system. Concluding remarks will be made in §7.

2. Abstractive and constructive models

Although there are many models of morphological description, recognizing two perspectives 'constructive' and 'abstractive' captures the fundamental differences in the ways all approaches hypothesize that surface inflected forms in a language are generated. The central tenets of constructive and abstractive models will be outlined here.

[1] The terms 'constructive' and 'abstractive' are from Blevins (2006).

[2] Although the term 'Word and Paradigm' is widely used when discussing abstractive approaches, it may, as has already been pointed out by Blevins et al. (to appear), be a somewhat misleading appellation since true abstractive approaches are unit-agnostic. The term 'Item and Pattern' captures this and is therefore adopted here.

In constructive models, inflected forms are conceived of as being 'built up' (constructed) from sub-word units, such as lexical stems and inflexional exponents. Morphemic models such as Item and Arrangement and Item and Process are two such approaches, as are 'root-based' or 'stem-based' perspectives, such as A-Morphous Morphology (Anderson 1992) and Paradigm Function Morphology (Stump 2001). Morphemic and stem-based models share the basic theoretical assumption that words are built up of recurrent sub-word partials, although they differ as to the size of the 'basic' unit of construction.

Firstly, in a morphemic account, morphemes are strictly concatenated onto one another according to language-specific patterns. In this type of approach, inflected word-forms are built in isolation, without reference to the paradigm. Speakers store lexical stems and meaningful sub-word elements, and this information, in addition to language-specific rules regarding the ordering of these units, is considered to be sufficient for the deduction of all inflected forms in a language.

In stem-based constructive accounts, on the other hand, wordforms are not built up of morphemes; instead, the association of a lexeme with certain syntactic functions licenses the selection of particular lexical stems in addition to the inflexional affixes which are relevant for the syntactic context. These are expressed in terms of 'rules'. In this way,

> word forms are constructed by 'spelling out' the properties associated with syntactic preterminals, so that words have no existence outside the syntactic context in which they are constructed.[3]

The many weaknesses of morphemic models in particular have long been noted. According to an Item and Arrangement approach, for instance, each and every syntactic function in a language should be encoded by an overt 'morph' (usually an affix); the function~form mapping must be direct. A brief examination of the natural languages of the world, however, provides myriad examples in which this mapping is obscured. For instance:

1. Zero morphs - in many inflected forms, certain functions may not be expressed by overt affixation. For instance, in the form *sheep* in English, the plural is not marked morphologically. A morphemic model may instead posit a 'zero-morph' *sheep-Ø* which encodes the 'plural' function despite not having phonological form.

2. Cumulative morphs - particularly in languages characterized by a lot of fusional structure, multiple functions are often simultaneously encoded by a single, indivisible morph.[4] So, for instance, there is no (principled) way to split the form PUELLAM (the accusative singular form of the Latin noun PUELLA 'girl') into a lexical stem and two invariant morphs encoding 'accusative' and 'singular'. Instead, the final *-m* may be analysed as a cumulative morph which simultaneously encoding both of these functions.

3. Extended exponence - in a form such as Ancient Greek *elelýkete* 'you had unfastened', 'the realization of aspect and voice confounds any attempt to establish a biunique property-formative correspondence',[5] as is demonstrated in the following table:

[3] Blevins (2006: 535).

[4] It is worth noting here that cumulative expression is not the same as polysemy. For example, in Latin, the form RĒGIBUS 'to/by the king' is polysemous as it expresses both dative plural and ablative plural functions. Cumulative morphs are those which simultaneously express different feature values, such as both 'dative' and 'plural', in the same syntactic environment (in this example, the suffix -IBUS).

[5] Blevins (2016: 51).

Table 1: Exponence relations in Ancient Greek *elelýkete*[6]

e	**le lý**	**k**	**e**	**te**
PAST	PERFECT		PAST INDICATIVE	2nd PLURAL
		ACTIVE		

In this example, both the past tense and the perfect aspect functions are expressed by multiple, and sometimes discontinuous, sub-word morphs.

The above examples are not exhaustive of the ways in which the function~form mapping can be disturbed, but they are all representative of the types of difficulty met by morphemic models.

Stem-based accounts do not run into exactly the same problems as morphemic accounts. However, one major drawback which is common to morphemic and stem-based models is the often arbitrary splitting of 'stem' and 'inflexional affix'. This will be discussed further in reference to Viitso's (2012) analysis of Livonian nominals in §4.

Abstractive approaches, by contrast, do not require fully-inflected forms to be 'built up' from sub-word units. Instead, '[t]he word is a more stable and solid focus of grammatical relations than the component morpheme by itself.'[7] It is fully inflected wordforms, rather than sub-word partials, which provide the speaker with information which can be exploited in the production and interpretation of novel inflected forms. It is recognized that even whole wordforms are abstractions, rather than persistent units, but it is suggested that they are abstractions that tend to provide greater certainty for speakers than sub-word elements.

This means, therefore, that abstractive approaches are UNIT-AGNOSTIC, which means that sub-word units, periphrastic constructions and idioms, for instance, may also be recognized and exploited by speakers in the production of novel forms. Stems and inflexional exponents can, therefore, be considered 'units of analysis' where they provide a speaker with predictive information, but they are not assigned meaning in isolation as in many constructive accounts.

In this sort of approach, there is no need for function and form to be in one-to-one correspondence (which is very rare in natural language). Instead, it is the implicational patterns that exist between fully inflected forms of nominals which are exploited by speakers in the production and interpretation of novel inflected forms. These relations are abstract and exist only between two or more related forms (whether those forms belong to the same lexeme, to the lexical neighbourhood or to the same word-class as another lexeme). These relations can be exploited by language users without the problems associated with assigning consistent morphological material to particular functions or with non-arbitrary splitting of the lexical stem and inflexional affix.

Importantly, it is not just the forms of words themselves that provide the speaker with essential information about the forms of other nominals. Instead, it is the wordform IN ADDITION TO THE FUNCTIONS IT ENCODES which provides the relevant information about inflexional patterns a nominal can follow. Consider, for instance, the following proportional analogies:

1. PUELLA : PUELLA 'girl'
MENSA : ? 'table'
? : MENSAE

[6] Adapted from Blevins (2016: 51), originally discussed by Matthews (1991: 173).
[7] Robins (1959: 128). Emphasis added.

But:

2. PUELLA : PUELLA ‘girl’
TEMPLA : ? ‘temple’
? : *TEMPLAE

In the first example, the Latin word PUELLA ‘girl’ is nominative singular and is characterized by a form ending in *-a*; the genitive singular form ends in *-ae*. MENSA ‘table’ exhibits a similar structure in that it ends in *-a*, and on consideration of this surface form, we may suggest that the unknown ‘?’ form is MENSAE - which is, here, the correct form. However, on consideration of form alone, the noun TEMPLA ‘temple’, which, like PUELLA has a form ending in *-a*, may lead us to suggest that the unknown form marked by ‘?’ is *TEMPLAE, a form not found in this lexeme’s inflexional paradigm. Instead, knowing (usually from context) that the form TEMPLA encodes nominative/accusative plural functions rather than nominative singular functions permits the speaker to determine that this lexeme does not follow the same inflexional pattern as PUELLA and MENSA.

Nevertheless, as mentioned above, sub-word units, where they provide some element of certainty for the speaker, are important in an Item and Pattern approach. Where recurrent partials occur and can be relied on by language users, I suggest that these may also be exploited in the production of novel forms and it is therefore essential to afford them a formalized place in an abstractive model of grammatical description. This will be discussed further in reference to Livonian in §5.

3. The Livonian language

Livonian is a Finnic language which was natively spoken in the north of Latvia until recent years.[8] The language began to be systematically documented relatively late - in the 19th and 20th centuries[9] - and died not long afterwards. Since Livonian is no longer natively spoken, the data presented here come from written sources. The basis of the written language is an eastern dialect, spoken in Courland in the west of Latvia.[10] The case system of the written variety will be discussed here.

Throughout, *ḑ, ţ, ļ, ŗ, ņ, ŗ, š* and *ž* are used for the palatalized consonants, *õ* is used to indicate the high central vowel,[11] *ȯ* is used for the mid-high back vowel and *ǭ* is used for the (long) mid-central vowel (with no short equivalent).[12] A macron distinguishes long vowels from short, whereas long consonants are written as geminates, according to convention. Finally, an apostrophe indicates the ‘broken tone’, ‘which is rising-falling or predominantly falling and articulated with laryngealization (stød or creaky voice)’.[13] The plain tone is not marked and occurs on initial syllables where the broken tone does not.

I will be focusing here on an analysis of the Livonian inflexional case system. Livonian has the fewest morphological cases of any Finnic language; grammars usually count

[8] The exact date of the death of the last speaker of Livonian is recorded differently by various sources. The Times offers 2013 (https://www.thetimes.co.uk/article/death-of-a-language-last-ever-speaker-of-livonian-passes-away-aged-103-8k0rlplv8xj), but Kittilä and Ylikoski (2011:48) record the date the death of the last speaker of Livonian as 2009.
[9] Grünthal (2015: 12).
[10] Pajusala (2014: 149f.).
[11] The precise phonetic value of particularly *õ* varies depending on the dialect. See Pajusala (2014: 154-161) for details.
[12] Viitso (2016: 146).
[13] Viitso (2007: 47).

between eight and eleven distinct (productive) forms.[14] The exemplary paradigm of *läpš* 'child' offers a typical example of the sorts of inflexional patterns exhibited by Livonian nouns and adjectives:

Table 2: *läpš* 'child' fully declined

	Singular	**Plural**
Nominative	*läpš*	*lapst*
Genitive	*laps*	*lapst*
Partitive	*lapsta*	*läpši*
Dative	*lapsõn*	*lapstõn*
Translative	*lapsõks*	*lapstkõks*
Elative	*lapsõst*	*läpšīst*
Inessive	*lapsõs*	*läpšīs*
Illative	*lapstõ*	*läpšīz*
Comitative	*lapsõks*	*lapstkõks*

The main functions of the Livonian case-forms are as follows:

- nominative: encodes the syntactic subject;
- genitive: used to express possession and the definite object;
- partitive: used to express the indefinite subject or indefinite object;
- dative: marks the 'recipient', 'beneficiary' and, sometimes, 'source';
- comitative-translative: encodes sociative, translative (predicative adverbial) and instrumental meanings;
- illative: 'towards' or 'into' (temporally and spatially);
- inessive: 'in' (temporally and spatially);
- elative: 'away from' or 'out of' (temporally and spatially);
- (instructive: 'expresses regulation of activity by items [...] or of amount by units or number',[15] e.g., 'I go there *by/on foot*', 'I attend church *on Sunday*s'. Lexically restricted);
- (essive: used to express a (temporary) state or function. Rarely encoded by a form distinct from the dative in the modern language);
- (abessive: expresses the caritive meaning 'without'. Lexically restricted).

For a more in-depth discussion on the functions the cases in Livonian, see Viitso (2016:147-151).

Throughout the language, the genitive singular form is rarely differentiated from the nominative singular form, and then almost always by morphophonological differences in the stem, not by overt affixation.

The Livonian inflexional case system is also characterized by sometimes extensive and systematic syncretism. In table 2, for instance, the nominative plural and genitive plural functions are encoded by identical forms (a pattern found in the inflexional paradigms of all nouns and adjectives), as are the comitative and translative functions (although distinct across number). Additionally, in many nominal types, the nominative singular and genitive singular (and, occasionally, the partitive singular) case-forms are syncretic. Where the essive is considered to be a productive case-form in Livonian, it is almost always encoded by the same

[14] The 'core' case-forms which are discussed are: nominative, genitive, partitive, dative, translative, elative, inessive, illative and comitative. Viitso (2016: 146) distinguishes separate essive and instructive forms; Grünthal (2003) omits the essive and instructive and considers the translative-comitative a single form encoding 'translative' and 'comitative' functions.

[15] Viitso (2016:148).

form as the dative.[16] In nouns and adjectives with a partitive singular form ending in *-Dõ* (not *-õ*), the illative singular form is distinct, but often these two forms are identical.

In written Livonian, the external local case-forms found in other Finnic varieties are 'now restricted to farmstead names, certain village names, place-names ending in *mǭ*, to a few nouns (and their adjectival attributes) [...] and to some adverbs and postpositions',[17] although there is no longer any morphological distinction between the adessive and allative even in these lexemes.[18] The meanings once encoded by these forms are now largely expressed by means of postpositions or by the new dative form.[19]

Finally, there are a few lexemes which differentiate the comitative and translative functions morphologically, e.g., *naizõks* 'for the wife' vs. *naizkõks* 'with the wife', *sūrõks pi'ņņõks* '[grows] into a big dog' vs. *sūr piņkõks* 'with a big dog'.[20]

The Livonian inflexional case system is characterized by extensive allomorphy, in both the lexical stem and the inflexional suffix, which cannot (always) be attributed to synchronically regular or automatic phonological alternations. Stem allomorphy will be discussed further in §4. The Livonian case suffixal allomorphs are listed in table 3:

Table 3: Livonian case suffix allomorphs[21]

	Singular	**Plural**
Nominative	*Ø*	*-d, -õd, -t,* (*-g, -nt, Ø*)[22]
Genitive	*Ø, -z*	*-d, -õd, -t,* (*-nt*)
Partitive	*-dā, -tā, -da, -ta, -ţa, -dõ, -tõ, -t, -õ, -i*	*-di, -ti, -ḍi, -ṭi, -i, -īdi, -idi*
Dative	*-n, -õn, -dõn*	*-dõn, -õdõn, -tõn*
(Ablative	*-ld(õ(st)), -õld(õ(st))*	*-iḷd*)
(Adessive	*-l(õ), -õl*	*-iḷ*)
(Allative	*-l(õ), -õl*	*-iḷ*)
Elative	*-stõ, -st, -õst, -št*	*-šti, -īst(i), -ist(i)*
Inessive	*-sõ, -s, -õs, -š(õ), -š*	*-ši, -īs(i), -is(i)*
Illative	*-zõ, -õ(z)*	*-ži, -īž, -iž, -ž, -īz, -iz*
Translative/comitative	*-kõks, -ks, -õks*	*-dkõks, -tkõks, -dõks, -õdõks, -tõks*
Instructive	*-īņ, -iņ*	

As already mentioned, the genitive singular is not usually marked by a distinct affix. However, in one nominal type, it is differentiated from the nominative singular by overt suffixation of an exponent *-z*.

To sum up:

> Livonian exhibits both complicated variation of inflectional suffixes and a complicated system of morphophonological alternations. In this respect Livonian exceeds, in the Finnic

[16] Viitso (2016) lists a few lexemes which distinguish which have distinct essive and dative forms.
[17] Viitso (2016: 147).
[18] Ernštreits and Kļava (2014: 80).
[19] Livonian is the only Finnic variety to have developed a dative form. For an account of the Livonian dative and its historical roots, see Inaba (2015). N.B. that the Salaca Livonian dialect does not exhibit a morphological dative form, and instead uses a productive adessive-allative form. Pajusalu (2014:162).
[20] Once again, the Salaca dialect differs slightly from the written variety in that the translative and comitative functions are consistently encoded by distinct forms.
[21] These are the allomorphs identified and listed by Viitso (2012: 20).
[22] The allomorphs in brackets are not found in nouns and adjectives.

> space, even the complexity of the considerably better studied Estonian.[23]

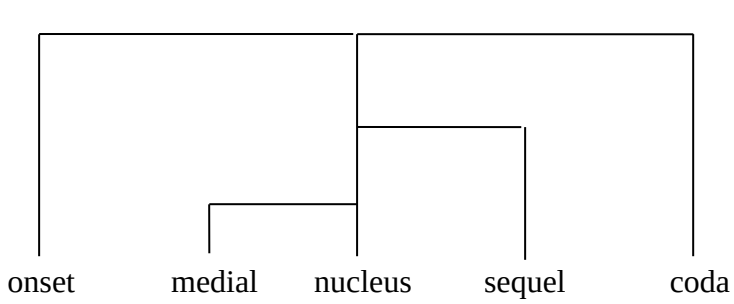
4. Previous analysis of Livonian nominals – Viitso (2012)

In his 2012 article, *Concerning inflection classes in Livonian*, Viitso outlines the different types of stem and suffixal allomorphy found in the inflexional paradigms of nominals and verbs. Although the 'ending variants and morphemic alternations served as a basis for establishing 62 verb types and 238 nominal types',[24] Viitso's stem-based analysis leads him to identify 24 classes for verbs and 85 classes for nominals. In the following discussion, I will consider only his analysis of nominals.

Viitso's analysis is constructive in nature, relying on stem allomorphy and its distribution within the paradigm as the means of differentiating nominal classes. He identifies three different types of morphophonological alternation exhibited by nominal stems: 1) that having to do with gradation; 2) that having to do with non-gradational tone alternation and 3) 'a heterogeneous set of non-gradational segment losses and alternations in word stems' which are distinct from those found in stem grade alternations.[25] Each of these will be discussed below.

Firstly, those nominals which exhibit gradation are characterized by two different stem grades: 'weak' and 'strong'. In order to understand the differences between weak and strong stem grades, a brief introduction to Livonian syllable structure is in order, since '[t]he system of quantity patterns in Livonian requires application of a more detailed syllable structure model than the traditional onset-nucleus-coda model'.[26] Viitso (2007) provides the following syllable structure for initial syllables:

Fig. 1: Structure of Livonian stressed syllables



> This model gives prominence to all positions where quantity contrast [sic] occur (medial, nucleus, and sequel) and separates them from the positions that do not participate in morphological alternations (onset and coda).[27]

With this structure in mind, Viitso offers the following distinctions between the two different stem types in Livonian: 'strong'-grade stems are characterized by 'a short nuclear vowel and a coda in the first stem syllable'; in 'weak'-grade stems on the other hand, the

[23] Viitso (2012:12).
[24] Viitso (2012: 12).
[25] Viitso (2012: 15).
[26] Viitso (2012: 14).
[27] Viitso (2012: 14).

coda is either absent or light ('which means that coda consonants and vowel are phonetically half-short or full-short').[28]

Viitso identifies four different types of gradation, most of which have a number of sub-classes. I will describe only the general grade alternation patterns here:[29]

1. disyllabic vocalic weak-grade stems alternate with monosyllabic strong-grade stems (e.g., *kalā-d* nominative/genitive plural 'fish', *ka'll-õ* partitive/illative singular);
2. disyllabic consonant weak-grade stems alternate with disyllabic vocalic strong-grade stems (e.g., *azūm* nominative singular 'place or situation', *a'zmõ-d* nominative/genitive plural);
3. a monosyllabic weak-grade stem in the partitive singular form alternates with a monosyllabic strong-grade stem in all the other forms (e.g., *laps* genitive singular 'child', *laps-tā* partitive singular);
4. a monosyllabic weak-grade consonant stem has a long nucleus and a short sequel and differs from the first syllable of the corresponding disyllabic strong-grade stem either in nucleus and sequel length or in nucleus length and the loss of the vocalic sequel (*võrgõ* nominative/genitive singular 'net', *vȭr-ta* partitive singular).

All nominals which exhibit grade alternations in their inflexional paradigms follow one of these patterns. For more information, see Viitso (2012).

The second type of alternation found in the stem is that of a non-gradational tone change. Two nominal types exhibit tone differences which do not have to do with grade alternations:

1. those monosyllabic forms with a long monophthong or a long opening diphthong exhibit an illative singular form characterized by the broken tone (e.g., *pā̄* 'head', illative singular form *pā̄'zõ*);
2. *mīez* 'man' has the level tone in only the nominative singular and partitive singular forms.

Finally, non-gradational alternations in word stems consist of the loss of segments (for instance, the syncope of certain vowels, the loss of *-z* and *-t* from stem-final position, etc.) as well as differences in the position of segments in different inflected word-forms (for instance, the metathesis of certain consonants, as in nominative singular *azūm* and nominative plural *a'zmõd*).

In addition to extensive stem allomorphy, Livonian is unique in the Finnic space in exhibiting suffixal allomorphs for ALL productive case values. The exception to this generalization is the nominative singular, which is always characterized by a LACK of suffixal case marker. The non-productive external local case-forms are invariant in the plural but elsewhere even the fossilized forms exhibit allomorphy.

Viitso identifies the following constraints on the co-occurrence of stem and suffixal forms that hold across the language:

1. where an inflexional suffix has allomorphs of the shape *-õC(C)* and *-C(C)õ*, the first occurs where the lexical stem ends in a consonant and the latter occurs with lexical stems ending in a vowel and 'optionally certain consonantal stems';[30]

[28] Viitso (2007: 48).

[29] N.B., in the following discussion a 'vocalic' stem is one which ends in a vowel before the inflexional case ending; a 'consonant' stem is one ending in a consonant. One of the apparent weaknesses of Viitso's analysis is the often seemingly arbitrary splitting of lexical stem and inflexional suffix, which will be addressed at the end of this section.

[30] Viitso (2012: 23). Note that this means there are NO phonological environments which consistently licence the introduction of one of these morphs over another. This constraint may therefore be better described as a 'generalization' rather than an absolute rule.

2. endings beginning in *-t* never follow *-t, -ţ, -ņ, -ļ, -ŗ*;
3. endings beginning in *-ţ* never follow *-t, -ţ, -s, -š, -b, -d, -ḍ, -g, -m, -n, -l, -r, -v, -j*;
4. endings beginning in *-k* never follow *-p, -k, -s, - š*;
5. endings beginning in *-s* never follow *-ţ, -s, -š, -ņ, -ļ, -ŗ*;
6. endings beginning in *-š* never follow *-p, -t, -k, s, -š, -m, -n, -l*;
7. endings beginning in *-b* or *-z* never follow *p, -t, -ţ, -k, -s, -š, -m*;
8. endings beginning in *-d* never follow *-p, -t, -ţ, -k, -s, -š, -d, -ḍ*.

According to observations of the co-occurrence constraints obeyed by inflexional stems and inflexional affixes, Viitso identifies 85 classes to which nominals can belong, as opposed to the 238 classes identified on a purely formal approach.

However, Viitso's analysis suffers from a couple of major weaknesses. The most problematic of these is the often arbitrary distinction of 'lexical stem' and 'inflexional affix'. For instance, consider the form *teiõst* 'from the louse'. The nominative/genitive singular form is *tei*. In certain declensions, the elative singular form is characterized by an underlying genitive singular stem (consider the form *lapstõ* in the paradigm for *läpš* in table 2). If this stem is found in this form, then the elative singular allomorph must be analysed as *-õst*. However, nominals in other declensions 'build' the elative singular form on then partitive singular form, which for this nominal is *teiõ*. There is, however, no principled means of determining whether the stem should be analysed as *tei* + elative singular *-õst* or *teiõ* + elative singular *-st* as both are valid on consideration of other inflexional patterns in Livonian.

Nevertheless, the constructive account does identify recurrent sub-word partials that exist in the language, many of which can be invariably associated with particular functions and which are regularly distributed. Recognition of these sub-word instances of isomorphic function~form mapping are, I suggest, important in understanding how language users produce and interpret novel forms, and should be afforded a formalized place in abstractive models. This will be discussed further in §6.

5. An abstractive approach to the Livonian inflexional case system

The weaknesses of Viitso's (2012) account of Livonian inflexion classes suggest that there are problems with a constructive account of this language. Certainly, an abstractive approach captures some of the striking implicational relations that exist between fully-inflected wordforms in a way that constructive accounts cannot. The following section will briefly outline some of the most powerful implicational relations that are found in the language.

There are two different types of patterns which are useful in the production of inflected wordforms in Livonian: 'structural' relations and 'implicational' relations. Structural considerations have to do with the phonological shapes of nominals that hold across lexemes within the same inflexion class; implicational patterns refer to abstract relations that hold between two or more fully inflected word-forms within the paradigm of a lexeme.[31]

The following structural patterns hold across the language:

1.

[31] A brief note is required on the different stems in Livonian. Firstly, the 'genitive singular stem' has the same shape as the genitive singular; the 'nominative/genitive plural' stem has, similarly, the same shape as the nominative/genitive plural; the 'partitive singular stem' is not found in all nominals, but where it occurs it ends in a final *-õ* (not *-tõ*) and it is the same shape as the partitive singular form.

i. the dative singular form is always characterized by a final *-n*, where the vowel *-õ-* will intervene where the inflexional stem preceding ends in any consonant but *-v-* or *-j-* (i.e., not glides);
ii. the dative plural form is always characterized by a final form *-õn* which occurs with the nominative/genitive plural stem;
iii. the nominative/genitive plural form is always characterized by a final *-D*, where *-t* follows a sibilant and *-d* occurs elsewhere;
iv. the translative-comitative form always ends in *-ks* or *-kõks*; the distribution of these forms is morphologically determined;
v. the partitive plural form always ends in *-(D)i*;
vi. where an inflexional case desinence is characterized by a 'mobile vowel' (for instance, the inessive *-õs/-sõ*), the alternation generally has to do with the phonological environment in which it occurs, where the vowel preceding the inflexional ending occurs with nominal stems ending in a consonant, and the vowel following the inflexional ending occurs with nominal stems ending in a vowel;[32]
vii. in the plural, the sibilants found in the internal local case forms are palatalized except where another palatalized sibilant precedes the inflexional ending;
viii. the long *-ī-* found in the plural internal local case-forms and the partitive plural forms of certain nominals alternates with the short *-i-* found in the same positions according to grade alternations (for instance, the nominative/genitive plural form *kuţād* '(of the) invitations' has a long *-ī-* in the partitive and internal local case forms in the plural since its stem vowel is long; *a'zmõd* '(of the) beds', on the other hand, as a short *-i-* because its stem vowel is also short).

The following implicational patterns hold across the language:

2.
i. the nominative plural and genitive plural functions are always encoded by the same form;
ii. where a partitive singular form ends in *-õ* (not *-Dõ*), the illative singular form will be identical;
iii. where the genitive singular form ends in a final *-õ* and the illative singular form is identical, the illative singular form can take an optional final *-z*;
iv. the elative form can always be deduced from the inessive form by the addition of a *-t-* immediately following the sibilant of the inessive form (this holds across both numbers);
v. similarly, the inessive form can always be deduced from the elative form by the omission of the *-t-* immediately following the *-s/-š* of the elative form (this holds across both numbers);
vi. the dative singular and translative-comitative singular forms are found with the genitive singular stem;
vii. the dative plural and translative-comitative plural forms are found with the nominative/genitive plural stem.

Consider the following paradigms as representative examples of some of the implicational relations identified above. Further patterns that are particular to these lexemes will be outlined below:

[32] Recall, however, that this is a generalization and not a hard-and-fast rule.

Table 4: Sample paradigms of Livonian nominals[33]

	Sing.	**Pl.**	**Sing.**	**Pl.**	**Sing.**	**Pl.**
Nominative	*lēba*	*lēbad*	*rikāz*	*rikkõd*	*nai*	*naizt*
Genitive	*lēba*	*lēbad*	*rikkõ*	*rikkõd*	*naiz*	*naizt*
Partitive	*leibõ*	*leibḍi*	*rikāzt*	*rikīdi*	*nāizta*	*naiži*
Dative	*lēban*	*lēbadõn*	*rikkõn*	*rikkõdõn*	*naizõn*	*naiztõn*
Translative-comitative	*lēbaks*	*lēbadõks*	*rikkõks*	*rikkõdõks*	*naizõks*	*naiztõks*
Illative	*leibõ*	*leibži*	*rikkõ(z)*	*rikīž*	*naizõ*	*naižiz*
Inessive	*lēbas*	*leibši*	*rikās*	*rikīs*	*naizõs*	*naižis*
Elative	*lēbast*	*leibšti*	*rikāst*	*rikīst*	*naizõst*	*naižist*
Gloss	‘bread’		‘rich’		‘woman’	

Each of these example nominals is from a different gradation ‘type’ identified by Viitso (*lēba* from type 1, *rikāz* from type 2, *nai* from type 3), where types and patterns of stem allomorphy determine the inflexion class to which a nominal belongs. Nevertheless, in an abstractive approach, many implicational patterns can be identified which exist between the inflected forms of lexemes belonging to those which Viitso lists different types (in addition to those described above which hold across the language). For instance:

- the partitive, inessive and elative forms in the plural are all characterized by the same stem shape;
- the forms of the dative, translative-comitative, elative and inessive in the singular are characterized by the genitive singular stem;
- where a speaker is familiar with the genitive singular and partitive singular forms, they know that a lexeme exhibits weak and strong grades;
- where a speaker is familiar with the genitive singular and partitive singular forms, they know for certain the shapes of both the strong and weak grade stems.

It is important to note here that an abstractive account does not need to posit ‘rules’ by which a language user might ‘deduce’ one stem shape from another is not. Instead, it is supposed that certain frequent forms are likely to be encountered and learned, and that these provide the relevant information about the stem shapes found in the inflexional paradigm of a given lexeme without the need to put forward complicated phonological processes by which one form may be deduced from another. In Livonian, the genitive singular and partitive singular forms are usually sufficiently informative of the type of stem allomorphy exhibited by a nominal (at least in the singular - knowledge of the nominative/genitive plural and partitive plural forms may be required to provide a language user with sufficient information to produce plural forms). In an Item and Pattern approach, it is the distribution of stem allomorphs and how much information that they provide for the language user which is important.

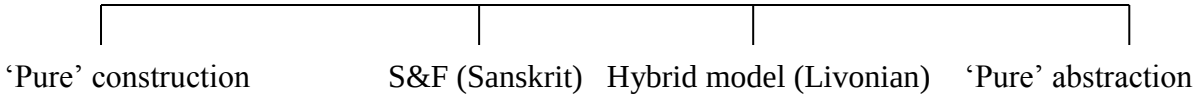
This offers a preliminary analysis of the structural and implicational patterns found in Livonian and does not pretend to be exhaustive. However, it will be noted that these descriptions have made use of both implicational relations (for instance: the distribution of stem grades) as well as sub-word units (such as the dative exponent *-n* and the elative element *-t-*). An abstractive account does not deny the importance of sub-word units, nor that they may provide language users with information about the production and/or interpretation of novel inflected forms. However, in recent work on this subject (most notably Blevins 2016), these sub-word elements are not offered a formalized place beyond playing a ‘discriminative’ role in language acquisition and use.

[33] Strong-grade stems are shaded.

6. Towards a hybrid constructive-abstractive model

I am not the first to suggest that elements from both constructive and abstractive models ought to be used in the description of the grammar of a language. Stump and Finkel (2013) have already argued for the benefits of a model which makes use of both the realization rules of a model such as Paradigm Function Morphology as well as the implicative relations found in, e.g., Blevins' (2016) Word and Paradigm approach. However, their model centralizes realization rules which are 'supplemented by one sort of implicative rule (rules of referral) to account for instances of syncretism.'[34]

Fig. 2. A gradient view of morphological models



In this gradient account of the models of grammatical description, Stump and Finkel's proposed model is closer to that of 'pure' construction rather than 'pure' abstraction.[35] My own approach, on the other hand, is closer to the 'pure' abstraction. That is, I believe that, in Livonian at least, implicational relations provide more relevant predictive information about the production and/or interpretation of novel forms than sub-word units and 'rules' regarding their combination. However, there are also a few instances in which greater predictive power is afforded by identifying sub-word recurrent partials and utilizing rules which determine their distribution in the language which ought to be reflected in the model with which the Livonian inflexional case system is described.

For instance, in Livonian, the dative, translative-comitative, elative and inessive across number and the nominative/genitive form in the plural in particular appear to lend themselves to hybrid interpretation. Whilst the stem shape found in these forms can differ depending on the lexeme, the inflexional endings are very often either invariant (e.g., the dative) or the allomorphs can be described as being distributed according to powerful and absolute phonological considerations (e.g., the nominative/genitive plural). Consider again the three lexemes in table 4. In certain instances a constructive account might better reflect the means that language users employ in the production (or interpretation) of novel forms. For instance, the dative ending *-n* is invariant across the language, and always occurs with the genitive singular stem in the singular and the nominative/genitive plural stem in the plural. When this form *-n* follows a vowel there are no changes; when it follows a stem ending in a consonant, an epenthetic *-õ-* occurs between the stem and the inflexional exponent (since the nominative/genitive plural form always ends in a consonant, the dative plural form may be better analysed as *-õn*).

A further example of an inflexional case form which can be accommodated in a 'rule-based' or constructive account is the nominative/genitive plural. In nouns and adjectives (although not all demonstratives or pronouns), the nominative/genitive plural form always ends in a dental *-d* or *-t*. These are distributed regularly according to the rules offered by Viitso, outlined in §4 and repeated here for ease:

1. endings beginning in *-t* never follow *-t, -ţ, -ņ, -ļ, -ŗ*;
2. endings beginning in *-d* never follow *-p, -t, -ţ, -k, -s, -š, -d, -ḍ*.

I also offered the following generalizations in §5:

[34] Stump and Finkel (2013: 294).
[35] For further detail, see their discussion in chapter nine, pages 271-294.

3. the nominative/genitive plural form is always characterized by a final *-D*, where *-t* follows a sibilant and *-d* occurs elsewhere.

The inter-predictability of the elative and inessive forms (discussed in §5) may also be better described in terms of 'rules' or construction more generally, or even in terms of both construction and abstraction, where these two are often difficult to disentangle. For instance, 'regular' plurals in English may be analysed as being produced by means of the addition of a final sibilant to a noun stem (e.g., *cat* + *s*), or by means of the analogous inflexional behaviour exhibited by other lexemes in the language (e.g., the implicational relationship of singular *hat* and plural *hats*). I suggest that, at least in certain instances, disentangling 'constructive' from 'abstraction' is at best difficult, if not impossible, which may provide further impetus to recognize the importance of, and develop, a hybrid theory of grammatical description.

Although language users may not store the distributional patterns described here as 'rules', knowledge of other forms and the distribution of the inflexional allomorphs throughout the Livonian inflexional case system provides information that can be used in the production and/or interpretation of novel inflected forms. In this way, 'rules' may be deduced through cumulated and extended exposure to the inflexional patterns in the language. More and more certainty is gained through further exposure to inflected tokens in the language, and the implicational patterns that can be identified may be exploited in language use. The language user employs whatever means are available in the production and interpretation of novel forms - particularly those means which provide the least uncertainty - whether this is constructive or abstractive means.

It will be noted that, since Livonian is no longer spoken natively, it is not possible to prove, with psycholinguistic experiments, for instance, whether the model proposed here accurately reflects the ways in which language users produce and interpret novel inflected forms. In order to gather conclusive evidence for the psychological reality of this hybrid model, therefore, a living language which is characterized by both agglutinative and fusional structure may be investigated. Estonian may prove a fruitful area of research.

7. Concluding remarks

A 'pure' constructive approach to morphological description does not adequately capture the generalizations that hold across the inflexional case system of Livonian, and does not permit a language user to correctly produce the expected inflected forms of a lexeme from the combination of sub-word lexical stems and inflexional exponents, whether by simple concatenation or by means of rules.

A 'pure' abstractive view, on the other hand, recognizes the importance of the inflexional patterns that exist between previously known inflected forms, and posits that these implicational relations are sufficiently informative in the production and interpretation of novel inflected forms. Despite the apparent complexity of the Livonian inflexional case system on account of its highly fusional structure (particularly when compared to the other Finnic languages), the patterns that hold between inflected forms are very powerful across the language. However, there are sub-word units in the Livonian language which are invariant, or which exhibit allomorphy which is distributed according to phonological considerations. The information that these provide for a language user may in some instances be better accommodated in a more constructive way. This has led me to suggest that a hybrid model of grammatical description may better accommodate the various inflexional patterns which are found in the Livonian morphological case system.

Current abstractive models have not yet clearly formulated a place for sub-word recurrent partials (although in a true abstractive account these can be analysed as 'information-bearing' where they can be consistently matched with particular functions). This paper has presented a 'hybrid' constructive-abstractive model in which sub-word units exhibiting an isomorphic mapping of function and form are afforded importance where they occur, but where it is not necessary to identify invariant forms and associate them with invariant functions in the production or interpretation of inflected wordforms.

References:

Anderson, S.R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Blevins, J.P. (2005). Word-based declensions in Estonian. In Booij, G. & van Marle, J. (eds.), *Yearbook of Morphology 2005*, 1-25. The Netherlands: Springer

----- (2006). Word-Based Morphology. *Journal of Linguistics* 42:3, 531-573.

Blevins, J.P., Milin, P. and Ramscar, M. (to appear). The Zipfian Paradigm Cell Filling Problem in Kiefer, F., Blevins, J.P. & Bartos, H. (eds.), *Perspectives on Morphological Structure: Data and Analyses*. Chapter 8. Leiden: Brill.

Ernštreits, V. & Kļava, G. (2014). Grammatical changes caused by contact between Livonian and Latvian. *Eesti ja some-ugri keeleteaduse ajakiri* 5:1, 77-90.

Grünthal, R. (2003). *Finnic adpositions and cases in change*. Helsinki: Suomalais-Ugrilainen Seura.

----- (2015). Livonian at the crossroads of language contacts. *Baltic Languages and White Nights*. Uralica Helsingiensia 7. Helsinki: Finno-Ugrian Society. 12-67.

Inaba, N. (2015). *Suomen datiivigenetiivin juuret vertailevan menetelmän valossa.* Helsinki: Suomalais-Ugrilainen Seura.

Kittilä, S. and Ylikoski J. (2011). Remarks on the coding of Goal, Recipient and Vicinal Goal in European Uralic in Kittilä, S., Västi K. and Ylikoski J. (eds.), *Case, Animacy and Semantic Roles*. 29-64 Amsterdam: John Benjamins Publishing Company.

Matthews, P. (1991). *Morphology*. Cambridge: Cambridge University Press.

Pajusala, K. (2014). The Relationship between Salaca Livonian and Courland Livonian Dialects. *Eesti ja soome-ugri keeleteaduse ajakiri* 5:1, 149-171.

Robins, R.H. (1959). In defence of WP. *Transactions of the Philological Society* 58:1, 116-144.

Stump, G. (2001). *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.

Stump, G. and Finkel, R.A. (2013). *Morphological Typology: From Word to Paradigm*. Cambridge: Cambridge University Press.

Viitso, T-R. (2007). Livonian Gradation: Types and Genesis. *Linguistica Uralica* 43, 45-62.

----- (2012). Concerning Inflection Classes in Livonian. *Linguistica Uralica* 48, 12-26.

----- (2016). The Essive in Livonian. *Eesti ja soome-ugri keeleteaduse ajakiri* 7:1, 145-175.

Zeprina-Jaz Ainsworth
Faculty of Linguistics, Philology and Phonetics
Clarendon Institute
Walton Street
Oxford
OX1 2HG
England

zeprina-jaz.ainsworth@worc.ox.ac.uk

ASPECTUAL VERBS: SIMPLE SEMANTICS + VARIABLE COMPOSITION

MICHAEL MOUROUNAS
(University College London)

GREGOR WILLIAMSON
(University College London)

Abstract

This paper discusses aspectual verbs and the core lexical semantics shared by their instantiations with various complement types. We provide an overview of the behaviours common to such constructions, and discuss the difficulties that prototypical analyses face when trying to account for more than non-finite complements. We propose a simple semantics for aspectual verbs built on the mereological notion of *parthood* (Champollion & Krifka, 2016) which – when combined with a fluid system of computation – can account for aspectual verbs in its various forms. The proposal has a number of direct benefits, including a principled explanation for the deficient temporal properties of aspectual verbs' clausal complements. We then extend the account to address instances in which aspectual verbs combine with non-eventive DPs. Finally, we compare our account to the similar account of Piñango & Deo (2016). We show that, on the basis of cross linguistic data, our account is to be preferred.

1. Introduction

Aspectual verbs such as *begin* are unique in their ability to combine with a wide variety of syntactic objects (e.g. infinitival clauses, noun phrases, etc.). Furthermore, the meaning and behaviour of these verbs is consistent across constructions, despite variation in complement-type. This paper offers a lexical semantics for aspectual predicates couched in the system of mereology (Champollion & Krifka 2016), and which draws inspiration from recent work by Piñango & Deo (2016).

We argue that the lexical semantics proposed herein can account for the behaviour of these verbs across complement types. To do so, we suggest that aspectual verbs can combine with their complement via different compositional rules. Specifically, we make use of *Functional Application* (Heim & Kratzer 1998) and *Restrict* (Chung & Ladusaw 2004) plus *Existential Closure* (Heim 1982).

In §2 we motivate our lexical entries by showing that the grammatical behaviour of aspectual verbs persists in both infinitival-complement constructions and nominal-complement constructions. In §3, we provide a brief overview of our assumptions regarding so-called 'complex event nominalisations' (c.f. Grimshaw 1990, Borer 2003), specifically regarding their semantic denotations. We then present our semantics in §4 and discuss the benefits and corollaries of our approach in §5. In §6, we extend the account to constructions with eventive subjects, specificational constructions, and constructions with non-eventive DPs. Finally, §7 compares our account to the recent proposal of Piñango & Deo (2016).

2. Motivating a single lexical entry

Aspectual control verbs are notoriously ambiguous between raising and control (Perlmutter 1970). When they are control verbs, aspectual predicates such as *begin* can be classified as

exhaustive control verbs (Landau 2000). Exhaustive control (henceforth, EC) is categorised as an obligatory one-to-one identity relation between overt controller and implicit controlee. As illustrated in (1), this one-to-one relation bans plural predicates (such as *gather* or *together*) from appearing in the embedded clause.

(1) a. *John began to gather at six.
b. John continued to go to the movies (*together).

There are a number of empirical correlates associated with EC (Landau 2000, 2004, *et seq*.), including the *tense mismatch* generalisation as stated in (2) and illustrated in (3).

(2) *Tense mismatch generalisation*
EC predicates do not permit conflicting temporal modifiers between their matrix clause and embedded clause.

(3) a. ***Yesterday***, John began to go to the movies **tomorrow**
b. ***Last week***, Mary continued to run the London marathon **next week**.

Another correlation that concerns us is the *overt embedded subjects*, as stated in (4) and illustrated in (5) (see also Grano 2015).

(4) *Overt embedded subjects generalisation*
EC predicates do not permit overt subjects in the embedded clause.

(5) a. *John began (for) **Bill** to go to the movies.
b. *Mary continued (for) **Fred** to run the London marathon.

We note that the above EC-related behaviours obtain wholesale into English event nominal arguments. For example, as in the clausal domain, plural-subject modifiers (e.g. *together*) are illicit in nominal complements to aspectual predicates:

(6) a. *John began the inspection of the factory **together**.
b. *Mary continued the examination of the patient **together**.

Furthermore, the *tense mismatch generalisation* also manifests in these constructions; nominal complements cannot contain a temporal modifier that conflicts with the matrix tense:

(7) a. ***Yesterday***, John began the inspection of the factory **tomorrow**.
b. ***Yesterday***, Mary continued the examination of the patient **tomorrow**.

Finally, the *overt embedded subjects generalisation* obtains in nominal complements; prenominal genitive arguments are barred from naming the Agent of the nominal event in these constructions:

(8) a. *John began **Bill's** inspection of the factory. (Agent = Bill)
b. *Mary continued **Fred's** examination of the patient. (Agent = Fred)

Crucially, these behaviours are otherwise available to English event nominalisations when not embedded under an aspectual control predicate:

(9) a. An inspection of the factory **together** would be lovely.
b. The inspection of the factory **tomorrow** may be delayed until **next week**.
c. John observed **Bill's** inspection of the factory. (Agent = Bill)

3. English event nominalisation

3.1. Theoretical assumptions

Thus far, we have focused on English *complex event nominalisations* (Grimshaw 1990). Their reference to events and projection of corresponding verb argument (*the inspection **of the factory***) has led to many variable analyses which often make use of different amounts of verbal syntactic architecture within the nominal itself (cf. Alexiadou 2002, Borer 2003, 2014, Harley 2009, among others).

Regardless of internal structure, we observe that these nominalisations have the syntactic distribution of DP arguments. For example, they participate in *one*-substitution (see 10a), may bind a reflexive anaphor (10b), and may themselves be arguments of predicative adjectives (10c).

(10) a. Dr. Millhouse performed **an inspection of a patient** and Dr. Sanchez performed **one** too.
b. [**The examination of that patient**]$_i$ essentially completed **itself**$_i$.
c. The examination of that patient was particularly **aggressive**.

Thus – assuming a denotation for eventive *the* in (11) – we propose the prototypical denotation in (12) for English event nominalisations (see Moulton 2014 for a similar analysis).

(11) $[\![\text{the}_{\textit{eventive}}]\!] = \lambda P_{\langle v,t\rangle}\colon \exists!e\,[P(e)].\ \iota e\,[P(e)]$

(12) $[\![\text{the examination of the patient}]\!] = \iota e\,[\text{examine}(e) \wedge \text{Theme}(e) = \text{patient}]$

3.2. A problem for standard accounts

Given that typical accounts of aspectual verb semantics are built specifically with clausal complements in mind, it is unsurprising that nominal complements do not naturally follow. For example, we give a (simplified) denotation for infinitive-selecting *begin* in (13), based on the semantics of the progressive (Condoravdi 2009).

(13) $[\![\text{begin}]\!]^w = \lambda P_{\langle e,st\rangle}.\lambda x.\lambda e.\ P(e) = 1 \text{ in } w \wedge \text{Agent}(e) = x \text{ in } w \wedge$
$\exists\langle e',w'\rangle[w' \in IN_w \wedge e \subset e' \wedge P(e') = 1 \text{ in } w' \wedge \text{Agent}(e') = x \text{ in } w']$
(adapted from Landau 2015)

Given the denotation of the event nominal in (12), the issue becomes clear: *begin* in (13) expects as its input a property of individuals (type ⟨e,st⟩), but our nominal complement denotes a unique event (type v).

Thus, we have a type mismatch, which leaves us with two options: (i) accept (13) as the lexical semantics for infinitive-taking *begin*, and propose a new and separate lexical entry for *begin* which specifically selects nominal complements, or (ii) reject (13), and propose a new lexical entry for *begin* which combines with both infinitive- and nominal-complements.

Since option (i) would render all of the shared behaviours noted in §2 an empirical accident, we take option (ii) to be the preferred choice.

4. The proposal

4.1. A radically simplified semantics for aspectual verbs

We propose that aspectual predicates select for *events*, rather than *properties*. Furthermore, we argue that these predicates encode a parthood relation $\leq$[36] between the selected event and the event denoted by the aspectual verb. To illustrate, we give a novel denotation for *begin* in (14). Note, we assume that the external argument is severed from the verbal semantics in the sense of Kratzer (1996).

(14) $[\![\text{begin}]\!] = \lambda e'\colon \exists e'' < e'.\lambda e.\ e \leq_{initial} e'$
where $e \leq_{initial} e'$ iff $e \leq e' \wedge \forall e''\ [e'' \leq e' \wedge \neg (e \circ^{37} e'') \rightarrow \tau(e'') \succ \tau(e)]$

The lexical entry in (14) states that *begin* will select an event which must be non-atomic. The subsequent parthood relation ensures that the *begin* event is the initial sub-event of the selected event by stating that the runtime of every other (non-overlapping) sub-event of the complement event occurs after that of the *begin* event.

By measuring event sub-parts by event-runtime ($\tau(e)$), we can define other aspectual predicates with minimal changes to the semantics in (14). For example, defining a lexical entry for *finish* is as simple as reversing the event-runtime precedence relation, and adding a presupposition that the event in question has already begun:

(15) $[\![\text{finish}]\!] = \lambda e'\colon \exists e'' < e'.\ \lambda e.\ e \leq_{final} e'$
presupposes $\exists e''\ [e'' \leq_{initial} e' \wedge e \neq e'']$
where $e \leq_{final} e'$ iff $e \leq e' \wedge \forall e''\ [e'' \leq e' \wedge \neg (e \circ e'') \rightarrow \tau(e) \succ \tau(e'')]$

Non-ordinal aspectual predicates – such as *continue* – are simpler still to characterise; since they assert an unspecified parthood relation relative to event-runtime, we require only the presupposition that the event in question has begun:

(16) $[\![\text{continue}]\!] = \lambda e'\colon \exists e'' < e'.\ \lambda e.\ e \leq e'$
presupposes $\exists e''\ [e'' \leq_{initial} e' \wedge e \neq e'']$

4.2. Variable methods of composition

With our new semantic denotations in hand, composition of aspectual verb with event nominals may now proceed straightforwardly via *functional application*:

(17) a. $[\![\text{the examination of the patient}]\!] = \iota e'\ [\text{examine}(e') \wedge \text{Theme}(e') = \text{patient}]$
b. $[\![\text{begin the examination of the patient}]\!] = \lambda e.\ e \leq_{initial} \iota e'\ [\text{examine}(e') \wedge \text{Theme}(e') = \text{patient}]$

[36] Def. as: $\forall x, y \in U\ [x \leq y \rightarrow x \oplus y = y\ ]$, where $\oplus$ is the two-place *sum operation* (Piñango & Deo 2016)
[37] The *overlap* relation, def. as: $\forall x, y \in U\ [x \circ y \leftrightarrow \exists z \in U\ [z \leq x \wedge z \leq y\ ]]$ (Piñango & Deo 2016)

In order to compose our aspectual predicates with infinitival complements, we employ the compositional method *Restrict* (Chung & Ladusaw 2004), which we define here for events.

(18) *Restrict*
Let γ be a node, and {α, β} the set of its daughters such that ⟦α⟧ is of type ⟨v,vt⟩ and ⟦β⟧ is of type and ⟨v,t⟩. Then, ⟦γ⟧ = λe.λe'. [⟦α⟧(e)(e') ∧ ⟦β⟧(e)].

We furthermore appeal to the notion that EC infinitives are relatively 'small', compared to other infinitival clauses; they are *v*Ps/VPs (Wurmbrand 2001, *et seq*.). As maximally *v*Ps, non-finite complements to aspectual verbs constitute *properties of events* (of type ⟨v,t⟩).

EC infinitives may therefore combine with aspectual predicates via *Restrict*. As an illustration, the infinitive *to examine patient* in (19a) restricts the open event position of *begin* to the sub-domain of events which are examining events, and whose Theme is *the patient*. An implicit step of *existential closure* in (19b) will yield a denotation very close to that in (17b).

(19) ⟦begin to examine the patient⟧
a. = λe'.λe. e $\leq_{initial}$ e' ∧ examine(e') ∧ Theme(e') = the patient
b. = λe. ∃e' [e $\leq_{initial}$ e' ∧ examine(e') ∧ Theme(e') = the patient]

5. Corollaries

There are a number of immediate benefits associated with our proposal. First, *exhaustive control* is entailed: since (e.g.) *begin* picks out a sub-event of its selected event, it follows that the subject of *begin* (introduced by Voice*/v)* is exhaustively a participant of (at least some of) the complement event.

(20) ⟦Mary begin to examine the patient⟧
= λe. ∃e' [e $\leq_{initial}$ e' ∧ examine(e') ∧ Theme(e') = patient ∧ Agent(e) = Mary]

Furthermore, our analysis makes some welcome predictions concerning aspect in the complement clause. We assume a fairly typical semantics for the perfective aspect (following Pancheva & von Stechow 2004), which asserts that the event time is properly contained within the reference time. The perfective aspect thus conveys that the event is completed.

(21) $⟦\text{PERFECTIVE}⟧ = \lambda P_{\langle v,t\rangle}.\lambda t_i.\ \exists e[P(e) \wedge \tau(e) \subset t\ ]$

The output of (21) will be a *property of times* (type ⟨i,t⟩), as such the denotation of AspectP would be unable to combine with the aspectual verb via *Restrict* due to a type mismatch. We thus predict the impossibility of temporal operators within the selected infinitive which occur higher than AspectP (i.e. progressive *be+ing*, and perfect *have*). This prediction is borne out:

(22) a. *John began to have been running the London marathon.
b. *Mary continued to have examined the patient.

Moreover, since the embedded infinitive cannot feature perfective aspect, we predict that there should be no entailment that the embedded event was completed. The prediction is also borne out, and best illustrated with telic predicates such as accomplishments (Vendler 1957), as in (23). Note, rather than having to appeal to the intentional semantics of the progressive (e.g. Condoravdi 2009), this behaviour falls out of our system for free.

(23) a. John began to climb the mountain.
b. ⇏ John climbed the mountain.

We likewise derive the *tense mismatch* generalisation without further stipulation. Note that we assume that temporal modifiers restrict the reference time, as in (24). Given the reduced nature of the infinitive in these constructions, we predict that the only time interval available for temporal modification will be that of the matrix clause. Thus, example (25) is ungrammatical because the two temporal modifiers are forced to modify the same time interval (i.e., the matrix reference time), leading to a contradiction.

(24) $[\![\text{yesterday}]\!] = \lambda P_{\langle i,t\rangle}.\lambda t.\ P(t) \wedge t \subseteq \text{yesterday}$

(25) a. *Yesterday, John began to run tomorrow
b. $\exists e \exists e'[e \leq_{\text{initial}} e' \wedge \text{run}(e') \wedge \text{Agent}(e) = \text{John} \wedge \tau(e) \subset t \wedge t \subseteq \text{yesterday} \wedge t \subseteq \text{tomorrow}]$

6. Extending the analysis

6.1. Eventive subjects

In many languages, we find constructions in which an eventive DP can act as the subject of a transitive aspectual construction. The example in (27) is from Brazilian Portuguese, and the example in (28) is from Mandarin Chinese.[38,39]

(26) The assassination of Franz Ferdinand started the first world war.

(27) A disputa começou a guerra
the dispute began the war
'the dispute began the war.'

(28) Guǎngdǎo hōngzhà jiéshùle dì èr cì shìjiè dàzhàn
Hiroshima bombing end-LE CARD 2 time world war
'The bombing of Hiroshima ended the second world war.'

On our account, this can be accommodated by assuming that an event can be construed as the external argument of an initial sub event. Namely, as its Cause.

6.2. Specificational constructions

Now, compare the causative eventive subjects in the previous section to the eventive DPs in the following *with*-PP constructions.

(29) a. The investigation began with the inspection of the factory.
b. The unveiling of the new building began with the cutting of the ribbon.

[38] Thanks to Natália Sousa (pc) for the Brazilian Portuguese data.
[39] Thanks to Ruoying Zhao (pc) for the Mandarin Chinese data.

In these sentences, the eventive DP introduced by the preposition is not construed as a Cause, but rather as a specification of the initial sub event. To see the difference, consider the following examples, given the context provided.

(30) Context: *The Arch Duke Franz Ferdinand was assassinated on June 28th 1914. After increasing tension, Austria-Hungary declared war on Serbia on July 28th.*
   a. The assassination of Franz Ferdinand started the first world war.
   b. #The first world war started with the assassination of Franz Ferdinand.
   c. The first world war started with the declaration of war by Austria-Hungary.

While the first of these sentences is true in the given scenario, the second is judged false. This is because the assassination of Franz Ferdinand was not a sub event of the first world war but rather a precursor. Compare this to (30c) in which the initial sub event of the first world war is specified. We propose to capture this difference by positing an unaccusative syntax for specificational constructions (31) and a non-instrumental preposition *with* that identifies events (32).

(31) $[_{VP}\ [_{VP}$ begin $[_{DP}$ the investigation $]\ ]\ [_{PP}$ with the inspection of the factory $]\ ]$

(32) a. $[\![\text{with}_{spec}]\!] = \lambda P_{\langle v,t\rangle}.\lambda e.\lambda e'.\ P(e') \wedge e' = e$
   b. $[\![\text{VP}]\!] = \lambda e.\ e \leq_{initial} \iota e'[\ \text{investigate}(e')] \wedge e = \iota e''[\text{inspect}(e'') \wedge \text{Theme}(e'') = \text{the factory}]$

6.3. Individual denoting DPs

Besides eventive complements, aspectual verbs can occur with non-event denoting DPs in specificational constructions.

(33) a. The row of toys began with a doll.
   b. The alphabet starts with the letter ‘A’.

However, these sentences become notably marked when they occur in a simple transitive construction.[40]

(34) a. ?? A doll began the row of toys.
   b. ?? The letter ‘A’ starts the alphabet.

Once again, this pattern seems to hold for languages as diverse as Brazilian Portuguese, Mandarin Chinese, and Saudi Arabic: non eventive DPs can only occur in specificational constructions with the aid of an adposition, and not in transitive constructions. The following minimal pair is from Saudi Arabic.[41]

(35) a. ṭābūr al’al’āb bādī ba’rūsa
   row the.toys start with.doll
   ‘The row started with a doll.’

[40] While sentences such as these are attested in archaic and higher registers of English, all the speakers we have consulted have found them significantly degraded.
[41] Thanks to Yara Al-Shaalan (pc) for the Saudi Arabic data.

b. *‘arūsa bada’t ṭābūr al’al’āb
doll start.F row the.toys
‘The doll started the row of toys.’

We account for this by treating individuals and events as sorted sub-domains of the same domain, D_e (Lasersohn 1995; Elliot 2016). Aspectual verbs then pick out the initial part of an entity according to some contextually supplied measure function μ.

(36) a. $[\![\text{begin}]\!] = \lambda y: \exists z{<}y.\ \lambda x.\ x \leq_{\text{initial}} y$
b. where $x \leq_{initial} y$ iff $x \leq y \wedge \forall z\ [z \leq y \wedge \neg (x \circ z) \rightarrow \mu(z) \succ \mu(x)]$

We suggest that the asymmetry noted above follows from the fact that, while an initial sub part of an entity can be specified (e.g., a doll is the initial sub part of the row of toys), there is no sense in which (e.g.,) a doll can be construed as the external argument of the initial sub part of the row of toys. That is, an event can be construed as the thematic Cause of another event, but an individual cannot be construed as a Cause of another individual, presumably because thematic functions are only defined for events. We thus find non-eventive DP complements in specificational constructions, but not causative constructions.

Consider finally complement coercion constructions with non-eventive DP complements. Our analysis predicts these constructions to be highly marked/unacceptable in many languages. In English, non-eventive DPs complements which do not stand in a part-whole relation with a DP in a specificational *with*-PP require a coercion operation to derive a suitable eventive reading. Thus, the utterance in (37a) can yield a number of possible interpretations (37b,c).

(37) a. John began the book.
b. → John began to read the book.
c. → John began to write the book.

These types of coercion constructions are particular to a subset of languages, and are not attested in others. By hypothesis, languages such as Czech and Mandarin lack a comparable coercion operator. These languages bar non-eventive DP complements in aspectual verb constructions, as illustrated in (38a). Event-denoting DPs are otherwise permitted in these constructions in Czech (38b).[42]

(38) a. *Jan začal knihu
John started the.book
‘John began the book’
b. Jan začal cesta
John started the.journey
‘John began the journey’

On our account, this follows straightforwardly, as there is no grammatical means by which the subject *‘John’* can constitute the external argument of the initial sub part of ‘*the book*’.

7. Comparison with Piñango & Deo (2016)

[42] Thanks to Zdeni Červená (pc) for the Czech data.

The semantics we have developed above is similar in spirit to that of recent work by Piñango & Deo (2016). However, there are crucial difference which render our account more constrained, and better suited to capture the cross-linguistic data. Below is P&D's type-flexible denotation for *begin*.

(39) $[\![\text{begin}_{P\&D}]\!] = \lambda y_{\sigma}.\lambda x_{\tau}.\ \exists f\,[f'(y) <_{initial} f(x)]$

The motivation for P&D's denotation is to provide a means by which to capture coercion constructions of the type noted in (37). However, we have suggested that, due to the cross-linguistic rarity of these constructions, we are better suited to keeping our semantics simple, while accounting for coercion constructions using additional implicit machinery (i.e., a coercion operator). Moreover, our semantics draws a clear distinction between specificational constructions on the one hand and transitive constructions on the other; the former is acceptable with both eventive and non-eventive DPs cross-linguistically, while the latter is only acceptable with eventive DPs, receiving a causative construal.

Finally, we have provided an explicit syntax-semantics which permits aspectual verbs to combine with the relevant complement types. This was done using independently proposed semantics for deverbal nouns, and independently motivated rules of composition.

8. Conclusion and outlook

We have provided a simplified lexical semantics for aspectual predicates which, when combined with a dynamic computational system, can account for the behaviour of these verbs across complement types. The motivation for such an approach came from the common behaviour displayed by aspectual verbs across domains; we focused primarily on infinitival and event-nominal complements. The lexical entries provided were based on the mereological notion of parthood and provide a schema upon which most aspectual verb semantics can be built.

Our proposal had a number of desirable consequences. The observable properties of the EC infinitive related to Tense and Aspect (e.g. the *tense mismatch generalisation*) followed naturally from our account. Furthermore, our assumptions concerning the nature of the complement made cross-linguistic predictions that were borne out.

The analysis provided in this paper draws some comparisons with that of Piñango & Deo (2016), whose account also leverages the mereological notion of *parthood*. However, our account is more constrained, and couched in a tangible syntactic system which permits principled composition with both VP and DP complements.

References

Alexiadou, A. 2002. *Functional Structure in Nominals: Nominalizations and Ergativity*. Berlin: Mouton de Gruyter.

Borer, H. 2003. Exo-skeletal vs. endo-skeletal explanations: Syntactic projections and the lexicon. In M. Polinksy & J. Moore (eds.), *The Nature of Explanation*. Chicago: Chicago University Press.

Borer, H. 2014. Derived nominals and the domain of content. *Lingua*, *141*, 71-96.

Champollion, L. & M. Krifka. 2016. Mereology. *Cambridge Handbook of Semantics*. Cambridge University Press.

Chung, S. & W. A. Ladusaw. 2004. *Restriction and Saturation*. Cambridge, MA: MIT Press.

Condoravdi, C. 2009. Measurement and instensionality in the semantics of the progressive. Ms., Stanford University.

Elliot, P. 2016. Explaining DPs vs. CPs without syntax. *ms.* UCL.

Grano, T. 2015. *Control and Restructuring*. Oxford: Oxford University Press.

Grimshaw, J. 1990. *Argument Structure*. Cambridge, MA: MIT Press.

Harley, H. 2009. The morphology of nominalizations and the syntax of vP. In M. Rathert & A. Giannakidou (eds.), *DP, QP and Nominalizations*. Oxford: Oxford University Press.

Heim, I. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. diss., University of Massachusetts, Amherst. Published 1987, NY: Garland Press.

Heim, I. & A. Kratzer. 1998. *Semantics in Generative Grammer*. Oxford: Blackwell.

Kratzer, A. 1996. Severing the verb from its external argument. In J. Rooryck & L. Zaring (eds.), *Phrase Structure and the Lexicon*, 109-137. Dordrecht: Kluwer.

Landau, I. 2000. *Elements of Control: Structure and Meaning in Infinitival Constructions*. Dordrecht: Kluwer.

Landau, I. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory*, 22, 811-877.

Landau, I. 2015. *A Two-Tiered Theory of Control*. Cambridge, MA: MIT Press.

Lasersohn, P. 1995. *Plurality, Conjunction and Events.* Springer Netherlands

Moulton, K. 2014. Simple event nominalizations: Roots and their interpretation. In I. Paul (ed.), *Cross-Linguistic Investigations of Nominalization Patterns*, 119-144. Amsterdam/Philadelphia: John Benjamins.

Pancheva, R. & A. von Stechow. 2004. On the present perfect puzzle. *PROCEEDINGS-NELS, 34*(2), 469-484.

Perlmutter, D. 1970. The two verbs *begin*. In P. Rosenbaum & R. Jacobs (eds.), *Readings in English Transformational Grammer*, 107-119. Waltham, MA: Ginn.

Piñango, M. M. & A. Deo. 2016. Reanalyzing the complement coercion effect through a generalized lexical semantics for aspectual verbs. *Journal of Semantics, 33*(2), 359-408.

Vendler, Z. 1957. Verbs and times. *Philosophical Review, 56*, 143-160.

Wurmbrand, S. 2001. *Infinitives: Restructuring and Clause Structure*. Berlin: Mouton de Guyter.