International Conference on Teaching and Learning English as an Additional Language, GlobELT 2016, 14-17 April 2016, Antalya, Turkey

L2 English Academic Speaking Development: Insights from a Multilingual University Context

Nicole Baumgarten[*]

University of Sheffield, 1 Upper Hanover Street, Sheffield S3 7RA, UK

Abstract

Little is known about the development of second language (L2) capacities in L2 users located in multilingual environments where more than one language is a viable communication tool and users can decide which to use for which purpose. Adopting a socially-grounded perspective on L2 learning, this study explores L2 academic English development in a multilingual university context in Denmark through a longitudinal study of 10 students' L2 performance in the academic register 'oral presentation'. L2 performance data were sampled on three naturally occurring classroom occasions during the students' first, second and final year of undergraduate study. The presentations were analysed for students' use of recurrent multiword sequences as a measure of development of routinized discourse production. This analysis was complemented by an analysis of the students' language use habits and socialization patterns. The analyses revealed some positive L2 English development between first and second but stagnant development between second and final year L2 performance for the investigated categories (tokens, types, structure, discourse function). This language use pattern coincided with students' orientation towards L1-based socializing by the final year. These results call for a reconsideration of academic L2 English instruction in multilingual environments outside native English-speaking settings, where L2 development seems to be susceptible to L2 users' overall language use habits.

 * Corresponding author. Tel.: +44-114-222-0546.
 E-mail address: n.baumgarten@sheffield.ac.uk

doi:10.1016/j.sbspro.2016.10.039

1. Introduction

The paper investigates the longitudinal development of students' advanced foreign-language (L2) English academic speaking performance outside a native (L1) English speaking context by the example of a multilingual university setting in Denmark. This study complements an earlier cross-sectional study of students' L2 English academic speaking development (Baumgarten, 2014). Based on research on L2 socialization (Duff, 2007; Kaloscai, 2008) and advanced L2 learning (Ortega & Byrnes, 2008), both studies start from the assumption that in multilingual and multicultural environments L2 development is susceptible to endemic notions of linguistic norms and communicative conventions as well as L2 users' individual needs and identity choices in their learning and usage environment. Individuals' ideas of linguistic norms are less consistent in multilingual settings than in monolingual settings (Le Page & Tabouret-Keller, 1985). In particular for the use of L2 English as a lingua franca it has been argued that when there is only little or no exposure to L1 register norms these are also of real normative value for L2 users (Canagarajah, 2007; Seidlhofer, 2004; Kaloscai, 2008). This affects L2 users understanding of the relationship between linguistic expertise and identity choices in a given community and may influence L2 users' understanding of learning needs.

Multilingual university settings are becoming more frequent throughout mainland Europe as universities increasingly offer bi-, tri-lingual (often including English) and international (English-medium) study programmes, typically under polces of *internationalization*. The purpose of internationalization is double-pronged: One aim is to produce multilingual professionals able to sell their multiple language capacities and multicultural expertise in an internationalized marketplace; the second is to open up national education systems to foreign students. The types of multilingualism produced and sustained by internationalized university settings is elite multilingualism, i.e. elective, voluntary L2 learning and use, adopted and maintained because it is believed to bring the individual educational and socio-economic advantages.

In internationalized university settings multiple languages are present which potentially constitute viable communication tools for people in and around that setting. When this multilingual space is located outside a L1 English-speaking context where English has no official status and role in the larger community, the use of English acquires a distinctive feature because it takes place in a space of reduced common ground. Common ground is here taken to be "presupposed common ground" (Stalnaker, 2002: 701), i.e. shared knowledge and expectations about, e.g., mutual access to local cultural knowledge and practices, language proficiency levels, English usage patterns, and communicative conventions and practices that bear on speaker and listener communicative behavior and that exist prior to the interaction. Fig. 1 illustrates the different degrees of common ground that can exist in internationalized, multilingual environments where English is used as a lingua franca.

Degree of mutually accessible common ground

- Expectations about local cultural knowledge and practices
- Expectations about English usage patterns and proficiency levels
- Expectations about shared communicative conventions
- Knowledge from ESL/EFL instruction

1. L2-L1 interaction inside a L1 English-speaking context
2. L2-L2 interaction inside a L1 English-speaking context
3. L2-L1 interaction outside a L1 English-speaking context
4. L2-L2 interaction outside a L1 English-speaking context
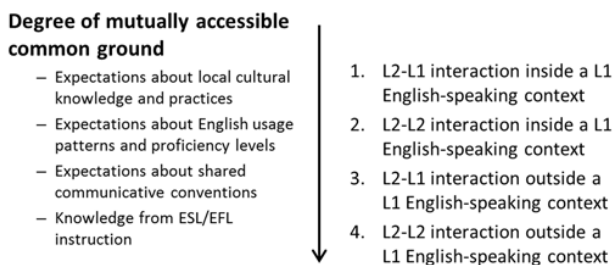
Fig. 1. Degrees of common ground

In constellation 4, reduced common ground can mean that there is no consistent norm-providing framework that regulates situational appropriateness of linguistic choice and, as a consequence, participants experience heightened variability in language forms, form-function mappings and interactional choice. It has been suggested that under

these circumstances normativity is suspended so that non-standard choice is not made interactionally relevant (Firth, 2009). The resultant unpredictability of linguistic choice opens up for L2 English users a multitude of L2 learning trajectories and socialization paths. When communicative constellations recur, the participants may evolve into a community of practice involving the acquisition and development of a linguistic repertoire of their own (Wenger 1998) through potentially idiosyncratic appropriation (Kramsch, 1998) of the L2. Accordingly, the research question for the present investigation was: How does students' L2 academic English evolve in the multilingual university space – when other languages may be viable means of communication in the environment and when L2 English use occurs mostly in L2-L2 interaction, i.e. under conditions of reduced common ground.

The study investigates recurrent multi-word sequences (RMS) as markers of L2 English linguistic and communicative development. Section 2 provides a description of RMS and their role in L2 development. Section 3 presents the language performance and self-report data used as well as the analytical procedures that were applied. Section 4 presents the major results of the analyses, and Section 5 concludes the paper with a discussion of the results in the light of the students' language learning and usage context.

2. Recurrent multi-word sequences (RMS)

RMS are the highly frequent sequences of words in a person's repertoire. Following Biber et al.'s (2004) analysis of "lexical bundles", high-frequency of multi-word sequences is taken to reflect pre-patterned status of the sequence so that it poses no production difficulty for the speaker. Fig. 2 shows an extract concordance for one RMS in the present data. 'this is' occurs 30 times and is used by 7 students. The sequence 'this is' is used in their talk as a stock phrase to identify a referent or a feature of a referent in discourse.

are mainly North America Australia and New Zealand? And this is a comparatively SMALL part of the world actually? I
can make mySELF and make OTHERS accept me to be. And this is a sentence from the book? And I really liked it
to talk about CROSS-CULtural BUSiness behavior. And this is actually the title of a BOOK which was written by n is
mostly done with REALLY DIFFICULT classical TEXTS. And this is also because this method was basic/ mostly used age
comes from withIN. Basically is a part of our DNA. And this is also connected with the concept of universal d ch
value placed on punctuality and on the scheduling? And this is for example the case in Germany? and there it is
cultures are more relaxed about time and scheduling. And this is for example the case in SPAIN? And conflicts can
times to cheating anyway. This is a very extreme you? And this is just to make it more easily to understand. To walk
CES across the NATIONS spoil the best efforts to sell. And this is not GOOD in the view of a BUSINESSMAN. So. Let's
very www/ just focused on on on on the deadlines. And this is not very good for the DEALS they wanted to close.
does does someb... Ya? Yeah. For example. Okay. And this is what I'm going to TALK about. I'm going to talk
it doesn't demand too much from from the READER? And this is why I I chose Gesteland's TEXT? But I also chose it
e p the person very well they are doing business with? And this is why it is very common in these deal-focused on the
picture. So everybody is EQUAL so to say? And this is why people from informal cultures address their we
are STUDYING INTERCULTURAL communication. And this is why we should at least be AWARE of these
is mostly due to luck. Or sometimes to cheating anyway. This is a very extreme you? And this is just to make it
uch more value on punctuality than OTHERS. But of course this is a stereotype. And you don't have to... This are not
definitions of language. I think THREE? But I found THIS is the best one? A systematic means of communicating
arises now? And the first is inherited cultural identities? This is identity that is FORMED by traditions stereotype and
He he also refers to his personal experiences. And ... Ja. This is like/ this is very helpful in order to to illustrate the

Fig. 2. RMS 'there is'

RMS occur in a speaker's repertoire as a function of time and exposure because speakers deal with recurrent communicative tasks in linguistically the same way in order to maximise communicative efficiency and minimise cognitive effort (Wray & Perkins, 2000). In this way, a pattern of linguistic choice for a particular communicative task becomes established. The quantitative presence of RMS in a speakers' speech reflects the degree of routinization of discourse production (Altenberg, 1998; Biber et al., 2004). RMS are used as building blocks of and scaffolding devices for continuous discourse, e.g. as open slot constructions (*this is why* + new information) or interlocking sequences ($_1$*I don't know if* $_2$*you know* $_3$*that this is* $_4$*one of the*). The more RMS speakers have in their repertoire and the longer they are the less ad-hoc online planning and speech monitoring is necessary. In addition, speakers experience themselves and will be experienced by listeners as more fluent (Graham & Barone, 2001),

which provides identity benefits in the speech situation. Differences in the frequency of RMS across time reflect a difference in the degree of routinization of discourse production and the discourse functions associated with the RMS. From a L2 developmental perspective, therefore, over time the use of RMS is assumed to be evolving. At later stages of L2 development, RMS have been described to be, e.g., more frequent, longer, and differently diversified (e.g. Adolphs & Durow, 2004; Wood, 2006). Reppen (2009) describes the longitudinal development of RMS as an indicator of register learning. None of the studies to date, however, have combined the characteristics of L2 English use in internationalized higher education, i.e. authentic, naturally occurring institutional talk, a setting outside a L1 English-speaking context, advanced L2 capacities as entry-level condition, and longitudinal data covering more than 12 months.

3. Data and method

We used a real-time longitudinal format to compare L2 performance during a three-year period. The data comprise classroom-based L2 performance data as an objective measure of L2 use and development and L2 users' self-report data as subjective assessments of their L2 learning and usage environment. The L2 users investigated are students in a trilingual 3-year undergraduate program in Modern Languages and Intercultural Communication at a small-size, regional university in Demark. They are Danish and German nationals with predominantly Danish and German L1 background. In this programme, students enrol in the first year with high-intermediate to low-advanced L2 English (equivalent to IELTS 6.5) and advanced beginner/low intermediate levels for L2 Danish and German (equivalent to 100-hours of instruction). L2 proficiency classes in Danish and German, and English LSP classes run throughout the program. The curriculum aims at balanced trilingual capacities at an advanced level with a focus on institutional communication. Beyond the language courses, the programme features English-medium teaching.

The study was designed to include the full group of students who enrolled in the programme's first year in the year 2011 (27 students). The final wave of data collection took place in 2013. Complete data sets exist for ten students. These ten classify themselves as either L1 speakers of German or Danish-German bilinguals.

The L2 performance data consist of three sets of student presentations from the first, second and third year of the programme. The presentations occurred as regular monologic speaking assignments in the LSP classes. We chose to focus this type of speaking because academic presentations have been described as high stakes communicative events for L2 speakers: They require L2 production over a prolonged period of time under real-time production constraints (Graham & Barone, 2001; Nesi & Basturkmen, 2009). The presentations were video-recorded and manually transcribed using an orthography-based transcription model (Rehbein et al., 2004). The sizes of the resultant YEAR1, YEAR2, and YEAR3 data sets are given in Table 1.

Table 1. Corpus information.

	YEAR1	YEAR2	YEAR3
Presentation topic	Free choice among topics of 'personal interest'	Free choice among topics of 'best business company'	Free choice among topics of 'most academically relevant concept'
Number of words	11721	15257	14350

The clusters/ngrams-function of the concordancer software *AntConc* (Anthony, 2012) was used to extract all RMS from the corpora. The cut-off frequency was set at three occurrences, i.e. to be counted as an RMS each sequence had to recur at least three times in any one presentation or across presentations. This procedure identified RMS shared by more than one speaker in each group as well as idiosyncratic language use. From the total of RMS those occurrences were eliminated which were specific to the topic of a presentation and recurred in only that presentation. The final sets of RMS were then analysed for their frequency distribution across years and speakers, their structural patterns, and discourse functions.

The self-report data were collected parallel to the recording of the presentations and targeted the students' socializing patterns in personal social networks and their language usage habits at the time of their presentational

speaking. The data were elicited in individual structured qualitative interviews conducted by student peers and from survey questionnaires administered by the researchers during class sessions.

4. Results

4.1. RMS

Tables 2 and 3 present the frequencies[†] of RMS types and tokens in YEAR1, YEAR2 and YEAR3 presentations. As a point of numerical comparison to academic presentational speaking by English L1 users, the tables include the frequency of RMS in the Arts and Humanities lectures component of the British BASE[‡] corpus and monologic discourse in the American MICASE[§] corpus.

Table 2. RMS types

	YEAR1	YEAR2	YEAR3	BASE_ah_lect	MICASE_monologic
2-word	48.7	48.6	55.1	57.5	56.2
3-word	9.0	12.0	13.3	32.2	36.6
4-word	2.0	2.3	2.2	9.1	10.3
5-word	0.3	0.3	0.3	2.2	2.1

The largest increase in RMS types occurs for 2-word RMS from YEAR2 to YEAR3. The number of 3-word RMS increases slightly from YEAR1 to YEAR2 and to a smaller degree also from YEAR2 to YEAR3. There is no development for the altogether infrequent 4- and 5-word RMS. (Larger RMS spans occur only at even smaller numbers and are excluded from this analysis). The frequency of 2-word RMS in YEAR3 reaches a level that is comparable to that in the L1 data; L2 speakers have considerably fewer RMS for all other spans, however.

Table 3. RMS tokens

	YEAR1	YEAR2	YEAR3	BASE_ah_lect	MICASE_monologic
2-word	280.8	312.6	342.6	597.8	654.8
3-word	37.1	53.4	55.1	189.4	222.0
4-word	6.6	7.9	7.6	41.8	46.3
5-word	1.2	1.0	1.0	8.5	8.7

RMS use (Table 3) increases for the shortest (2- and 3-word) RMS. 2-word RMS show a steady increase from YEAR1 to YEAR3; for 3-word RMS positive development is largely restricted to YEAR2. RMS use in L2 talk is far less frequent than in the L1 data. Overall, the RMS repertoire only grows minimally over time; a continuous increase in use is restricted to the shortest. The development of longer types and tokens peaks in YEAR2. For both RMS types and tokens, the difference between L1 and L2 use increases with RMS size.

Fig. 3 (below) shows the percentages of talk realized by RMS for each student across years. The L2 speakers use RMS in their talk to different degrees. For all speakers, there is – by YEAR3 – an increase in the portion of talk that is realized in a pre-patterned, routine fashion by utilizing RMS. This indicates that speakers learnt to perform linguistically in the presentation situation. One aspect of this learning is the repetitive use of word sequences in talk.

The 3-5-word RMS were categorised according to their structural types, following the typology revealed by Biber et al.'s (2004) analysis (Fig. 4 below). The categories are VP-based (*you can see on*), main clause with dependent

_____

[†] Normed frequencies on the basis of 1000 words are used throughout.
[‡] The British Academic Spoken English corpus project. Universities of Warwick and Reading.
[§] Michigan Corpus of Academic Spoken English, University of Michigan.

clause fragment (MCDC) (*I don't know if you*), dependent clause fragment (DC) (*which is the*), and phrasal fragment (*most of the*). Fig. 4 shows a very slight increase (approx. 1 per 1000 words) in 3-word RMS with MCDC and DC patterns by YEAR3. This means that speakers have acquired additional pre-patterned means to support complex clause construction in talk.
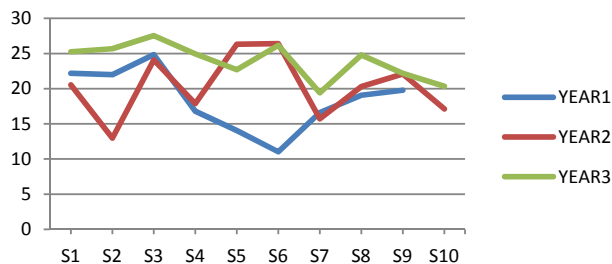

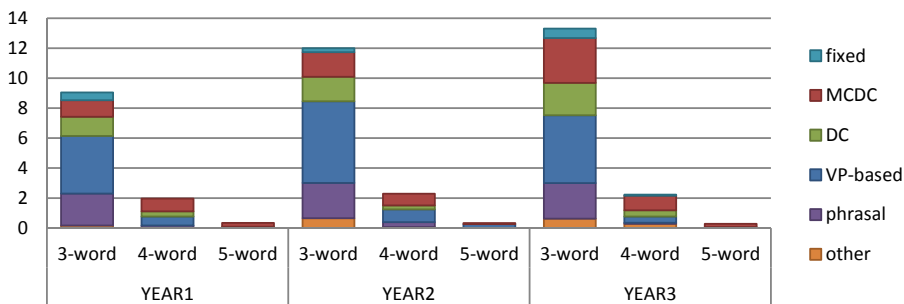Fig. 3. Percentage of RMS in individuals' talk.
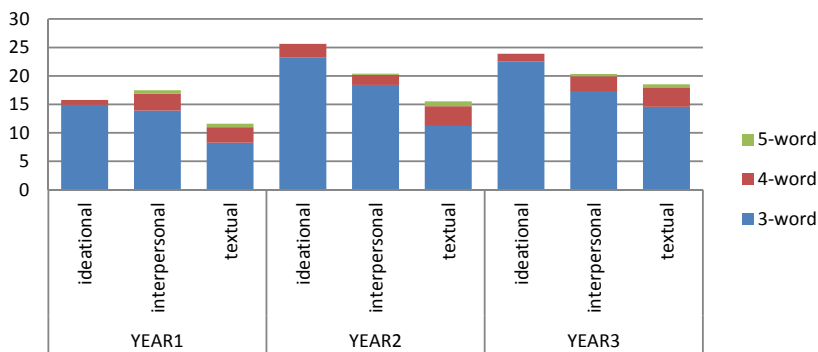

Fig. 4. Structural patterns of RMS types.

Fig. 5. Discourse functions of RMS tokens.

The 3-5word RMS tokens were also categorized according to the discourse functions they fulfil in their contexts of occurrence (Fig. 5 above). This categorization distinguished between referential, interpersonal and textual

functions, elaborated from the categories suggested by Biber et al. (2004) and Simpson-Vlach & Ellis (2010). This three-way categorization covers the three main tasks of presentational speaking, i.e. content presentation and topic specification, expression of speaker subjectivity and audience involvement, and multimodal discourse management.

By YEAR2, referential RMS are used most frequently, indicating the centrality of content-oriented talk in academic discourse, followed by interpersonal and textual RMS. From YEAR1 to YEAR2, there is an increase in the use of RMS for all three discourse functions, reflecting a routinization of carrying out these discursive tasks in the presentations. A further slight increase occurs for textual RMS from YEAR2 to YEAR3. The proportion of RMS sizes does not change over time, which indicates that either speakers do not acquire longer structures for completing discourse tasks or use interlocking to combine shorter structures into longer ones in individually unique ways.

4.2. Language usage patterns

Fig. 6 (below) presents the students language use habits inside and outside university on an average weekday. The figures were calculated from L2 users' responses to the interview question "Which languages do you use on an ordinary day inside and outside university?" Fig. 6 shows a drop in L2 English use from YEAR1 to YEAR3. At the same time German increases in importance.
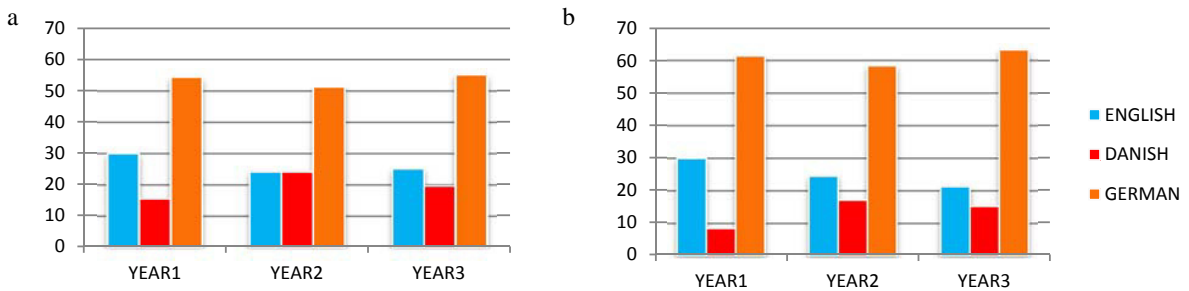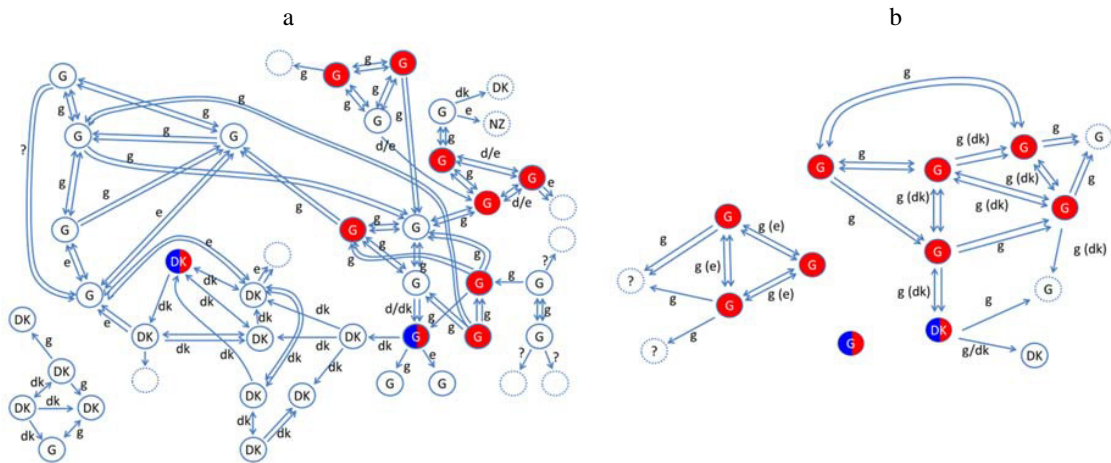

Fig. 6. Language use inside (a) and outside (b) university in percent.

Fig. 7 (below) represents the students' socializing patterns inside university. From YEAR1 on students form individual personal networks with fellow compatriots. In these networks the students L1 is the main means of communication.
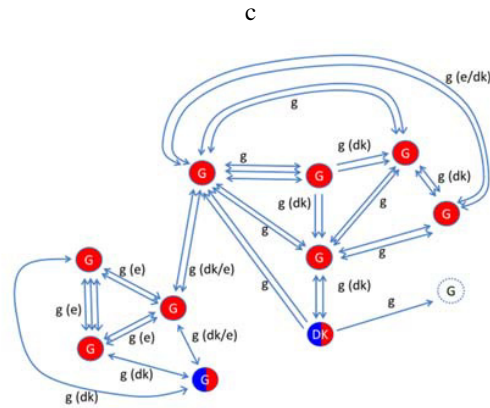
c

Fig. 7. Personal social networks YEAR1 (a), YEAR2 (b), YEAR3 (c). Coloured circles indicate the investigated students. Bi-coloured circles identify Danish-German bilinguals

By YEAR3, the networks have become closer and the relationships multidimensional (number of arrows) and reciprocal (two-sided arrows).

5. Conclusion

The comparison of the L2 production data revealed limited and uneven RMS development for the ten students under investigation. The size of their RMS repertoire increases, but predominantly in terms of the shortest, 2-word, sequences. Few longer tokens are added. The frequency of RMS use increases likewise mainly for short RMS; the increase for 3-word RMS is restricted to YEAR 2. There is a limited change in structural diversification toward more syntactically complex RMS. That RMS are experienced by speakers as useful in speaking is indicated, first, in the growing percentage of talk that is realized through RMS and, second, in the diversification across discourse functions. For the latter, however, development stagnates in YEAR2. Overall, the analysis suggests that RMS development and thus the automatization and routinization of discourse production peaks for this group of L2 users in their second year of studies in the multilingual university setting. From then on the importance of another language (German) increases in the students' lives, while English, toward the end of their studies, has become less relevant than at the start. This means that L2 learning, in terms of pre-patterning of language use with its attendant effects of better control of the speech situation, increased fluency, reduced processing effort, and indexing mature speaker identities does not occur simply as a factor of time and exposure in the multilingual setting. Speakers' L1 continues to exert substantial power as a pull factor in socializing and network building, effectively curtailing use and development of other languages in their repertoire – including English, although English is the dominant language of instruction.

Accordingly, one conclusion from these results is that agency in language choice can influence the path of development in the L2: Students' L1 remains an important and useful means of communication, and investment in L2 development is impeded by the subjective assessment of communicative needs in the multilingual space. From the first year on, the L1is important in relationship building and in selecting social ties, presenting an interpersonal condition which is difficult to change after relationships have been built. Another conclusion is that development in L2 academic English depends on a clear idea, first, what exactly academic English is in the international university setting outside a L1 English-speaking context, second, where it is supposed to exist in the multilingual university space, and, third, where educational institutions think students are supposed to get it from. Our analysis shows that the current combination of L2 English immersion in a multilingual setting plus L2 English instruction alone does not

support continuous learning. The final conclusion is that English-medium teaching and LSP instruction have to be re-thought as the main agents in L2 socialization if the objective is for students to develop functional capacities in English as an academic lingua franca beyond the immediate setting. This includes that the design of curricula and course contents must try to balance out students' multilingual repertoire to the degree that it is conducive to L2 academic English development. Academic English needs to be presented more clearly and more explicitly to students through more conscious deployment and modelling in English LSP instruction as well as in English-medium teaching if the aim is not to put L2 users in non-L1 English speaking settings at a disadvantage regarding the acquisition of academic communication skills.

References

Adolphs, S. and Durow, V. (2004). Social-cultural integration and the development of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 107-126), Amsterdam: John Benjamins.

Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp 101-122). Oxford: Oxford University Press.

Anthony, Laurence. (2012). AntConc.[AntConc 3.3.5w]. Available from http://www.antlab.sci.waseda.ac.jp/software.html.

Baumgarten, N. (2014). Recurrent multiword sequences in L2 English spoken academic discourse: Developmental perspectives on 1st and 3rd year undergraduate presentational speech. *Nordic Journal of English Studies* 13, 3, 1-32.

Biber, D., Conrad, S. & Cortes, V. (2004). *If you look at ...*: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25, 3*, 371-405.

Canagarajah, S. (2007). Lingua franca English, multilingual communities, and language acquisition. *The Modern Language Journal, 91, 1*, 923-939.

Duff, P. A. (2007). Second language socialization as sociocultural theory: Insights and issues. *Language Teaching, 40*, 309-319.

Firth, A. (2009). The lingua franca factor. *Intercultural Pragmatics, 6, 2*, 147-170.

Graham, J. G. & Barone, S. M. (2001). Academic speaking: Learning to take 'longer turns'. In J. Murphy & P. Byrd (Eds.), *Understanding the courses we teach: Local perspectives on English language teaching* (pp. 281-308). Ann Arbor: University of Michigan Press.

Kaloscai, K. (2008). Erasmus exchange students: A behind-the-scenes view into an ELF community of practice. *Apples – Journal of Applied Language Studies*, 3, 1, 25-49.

Kramsch, C. (1998). *Language and culture*. Oxford: Oxford University Press.

Nesi, H. & Basturkmen, H. (2009). Lexical bundles and discourse signalling in academic lectures. In J. Flowerdew & M. Mahlberg (Eds.), *Lexical cohesion and corpus linguistics* (pp. 23-44). Amsterdam: John Benjamins.

Ortega, L. & Byrnes, H. (2008). The longitudinal study of advanced L2 capacities. An introduction. In L. Ortega & H. Byrnes (Eds.), *The longitudinal study of advanced L2 capacities* (pp. 3-20). New York: Routledge.

Rehbein, J., Schmidt, T., Meyer, B., Watzke, F. & Herkenrath, A. (2004). Handbuch für das computergestützte Transkribieren nach HIAT. In *Arbeiten zur Mehrsprachigkeit*, Folge B 56, 1-90.

Reppen, R. (2009). Exploring L1 and L2 writing development through collocations: A corpus-based look. A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language* (pp. 49-59). Basingstoke: Palgrave Macmillan

Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics, 24*, 209-239.

Simpson-Vlach, R. & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics, 31, 4*, 487-512.

Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, 25: 701–721.

Wood, D. (2006). Uses and functions of formulaic sequences in second-language speech: An exploration of the foundations of fluency. *The Canadian Modern Language Review / La revue canadienne des langues vivantes, 63, 1*, 13-33.

Wray, A. & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication, 20*, 1-28.