Søren W. Eskildsen and Teresa Cadierno
University of Southern Denmark

Are recurring multi-word expressions really syntactic freezes? Second Language Acquisition from the perspective of Usage-Based Linguistics.* [1]

1. Introduction

This study focuses on the development in the oral use of negation patterns by a classroom Mexican learner of English. In theoretical agreement with recent claims made in various areas of second language acquisition (SLA) research, e.g., matters of frequency (Ellis 2002), learner constructions (Waara 2004), motion constructions (Cadierno 2004, Cadierno and Ruiz 2006), the aim of the study is to discuss the role of *recurring multi-word expressions* (MWEs) in L2 acquisition and use from the perspective of Usage-Based Linguistics (UBL). For our purposes UBL is particularly relevant as theoretical framework because it acknowledges the importance of MWEs and because it does not dichotomise syntax and lexis. This, we argue, allows for a better understanding of the structure, meaning, use, and acquisition of MWEs. In this paper we present our theoretical approach in more detail and present some results of our on-going investigations in SLA.

---

* The material presented in this paper is part of two ongoing research projects conducted by the two authors on the use of multi-word expressions by non-native speakers of English. One project, which is financed by a Ph.D.-grant at the University of Southern Denmark, takes ESL classroom interaction as its empirical point of departure. The other project, which is partly financed by the Danish Council of Humanities, focuses on the use of multi-word expressions in globalised organisational contexts, where English is used as a lingua franca.

[1] The authors would like to thank John Hellerman for useful comments on an earlier version of the paper.

2. Theoretical underpinnings

According to Tummers, Heylen, and Geeraerts (2005: 228), UBL comprises a number of related linguistic theories which are united by three main principles. These principles are "the priority assigned to language use, the integration of competence and performance, and the rejection of the rule-list fallacy." The first principle implies that usage is the only valid source for investigating the systemic nature of language. Language structure is, in other words, seen to emerge from concrete usage situations, and "it is the task of the linguist to study the whole range of repetition in discourse and to seek out those regularities that promise interest as incipient subsystems" (Hopper 1998: 166). In terms of the individual language user, this means that usage mirrors knowledge and that linguistic knowledge is conceptualised as linguistic experience (Tomasello 2000). It also implies a link between constructions as conventionalised in the speech community and as entrenched in the mind of the individual (Ellis 2002). This belief that language use and language knowledge, interaction and cognition, individuality and sociality are mutually constitutive is directly linked with the second UBL principle that competence and performance are seen as interwoven rather than dichotomous.

With respect to the third principle mentioned above, UBL rejects the view of language knowledge as dichotomised along the lines of 1) something that is rule-governed, cognitively demanding, and flexible and 2) something that is listed in the mental lexicon, swiftly processed, and stable. Related to the UBL conceptualisation of language knowledge as a structured inventory of symbolic units, i.e., conventionalised form-meaning pairings, used for communicative purposes (Langacker 1987), this is the rejection of the rule-list fallacy and its "[…] assumption, on grounds of simplicity, that particular statements (i.e., lists) must be excised from the grammar of a language if general statements (i.e., rules) that subsume them can be established (Langacker 1987: 29).

Adult language knowledge, then, is seen to consist of a continuum of linguistic constructions of different levels of complexity and abstraction comprising "concrete and particular items (as in words and idioms), more abstract classes of items (as in word classes and abstract constructions), or complex combinations of concrete and abstract pieces of language (such as mixed constructions)" (Tomasello 2003: 99). It should be kept in mind, however, that only *adult* language knowledge has this systemic nature. The emergent nature of language structure as described above means that ontogenetically, children do not learn their L1 on the basis of an innate Universal Grammar. Rather, operating "with different psycholinguistic units than adults" (Tomasello 2000: 62), children learn language in an item-based fashion heavily reliant on frequency, recurrence, and imitation.

Only when enough exemplars (items) have been encountered by the child does she begin to build the abstract cognitive schemas thought to underlie language knowledge. This view of language learning as item-based combined with the rejection of the rule/list fallacy means that storage and learning may take place on multiple levels simultaneously, involving both concrete items and abstract schemas, which may cohabitate in the grammar (Achard 2006). For example, storage for PLURAL may be dually represented as a) car + *s* = cars; and b) THING + [morph] = more things.

In opposition to the UBL principles outlined above, previous research into Formulaic Language (Wray's (2002) term; FL) has been predominantly 'syntactocentric', influenced by formalistic, generative views on language. This influence is clear in Wray (2002: 143) where it is stated that "language knowledge entails developing rules to generate all the possible utterances of the language". Dominated by the rules/list fallacy, this has lead to the view that whatever is 'formulaic' is somehow inferior to syntactic rules (Van Lancker-Sidtis and Rallon 2004). Applying the UBL framework, thereby challenging the strict division between creative expressions generated by syntax, on the one hand, and lexicon-based expressions on the other, we aim to challenge this. Whereas Wray (2002) finds it premature to implement UBL as a theoretical framework in research in FL, we argue that because of its attempts to account for all kinds of usage patterns without compartmentalising language, it allows for a better understanding of the role of MWEs in language usage and language learning.

3. Formulaic language in second language acquisition (SLA) research

With the publication of books by Wray (2002) and Nesselhauf (2005), the anthology edited by Schmitt (2004) as well as conferences such as *Collocations and Idioms: The First Nordic Conference on Syntactic Freezes* 2006, FL has come of age as a field of research. Even though Van Lancker-Sidtis & Rallon (2004: 211) describe the field as "handicapped by a bewildering array of variously defined terms", it is possible, in terms of SLA, to find some general trends in the existing research in FL. Without going into detail (but see e.g. Gitsaki 1999, Nesselhauf 2005, Wray 2002 for full discussions), we note that most research concerns collocational competence, Verb+NP, or Adv-Adj collocations, and is carried out with written language. Results are mixed in some respects, but the following conclusions are generally reached (Nesselhauf 2005): a) collocational production presents problems for learners, and more serious problems than general vocabulary use; b) learners use fewer collocations than native speakers; c) learners are not

aware of neither collocational restrictions nor combinatorial potential of lexical items.

Intriguingly, even though deviant collocational usage by L2 learners is sometimes reported to be as low as 25% (Nesselhauf 2005), statistics are nonetheless read as confirming the status of collocational competence as problematic for L2 learners. This, it is argued here, is a result of the underlying view FL as deviant from the norm of grammatically generated language. It is even implied in Wray (2002: 196) that "learning formulaic language is not 'real' language learning". Real language learning, in this view, presupposes analysability, combinability, and computation based on rules of syntax. In an inherently circular mode of argumentation, FL is seen as evidence that it is formulaic because it suggests inconsistencies between FL and grammatically combinatorial knowledge; learners are seen to produce language that is simply beyond their current interlanguage competence. In other words, FL is considered formulaic, i.e. *frozen*, because the underlying competence system, for some reason, could not have produced it – and this is so because it is formulaic.

From the perspective of UBL, the distinction presented above between FL and rule-governed syntax is a manifestation of the rule-list fallacy applied to SLA. Instead of making it a question of either FL or rules, UBL more conveniently accounts for the role of FL in language usage and language learning because it allows for dual representation, cf. above. In Tomasello's (2003: 106) words: "[…] in usage-based approaches a given linguistic structure may exist psychologically for the speaker both as a concrete expression of its own […] and at the same time, as an exemplar of some more abstract construction […] The main point from an acquisition point of view is that when a higher abstraction is made the lower level concrete constructions and expressions do not necessarily go away but remain available for use – especially if they are used frequently."

This introduces the notions of type and token frequency and their proposed importance for language acquisition. Token frequency refers to the idea that frequency of a concrete expression "in the language learner's experience tends to entrench that expression in terms of the concrete words and morphemes involved", whereas type frequency "of a class of expressions determines the abstractness or schematicity of the resulting construction" (Tomasello 2003: 107). In other words, storage as wholes is dependent on token frequency, whereas schematised knowledge and therefore productivity is dependent on type frequency (see also Ellis 2002).

With those aspects of frequency in mind, we set out to investigate for SLA the validity of the following proposed path of child language acquisition going from

concrete formulas via low-scope patterns (part concrete, part abstract) to abstract constructions (Ellis 2002, Tomasello 2000, 2003). Formulas are holophrases, single words, and so-called frozen phrases, sometimes idiosyncratic, which convey "a holistic, undifferentiated communicative intention, most often the same communicative intention as that of the adult expressions from which they were learned." (Tomasello 2003: 36). Examples are *lemme see, I wanna do it*. Low-scope patterns include pivot schemas, i.e., still concrete pieces of language, with one item structuring the utterance or determining the speech act, such as *where's the X, It's a X, X gone*, as well as verb-island constructions. These are utterance schemas revolving around specific verbs and constructions in which each verb is an island in the sense that morphological and syntactic markings are first learned on a verb-by-verb basis and not immediately generalised to other verbs. The final point on the proposed path concerns abstract constructions, i.e., higher-level constructional schemas that cut across different verbs, such as transitive constructions, negation constructions etc. The present study thus seeks to examine whether or not this "L1 acquisition sequence […] could serve well as a reasonable default in guiding the investigation in which exemplars and their type and token frequencies determine the second language acquisition of structure (Ellis 2002: 170).[2]

4. The study

4.1. Design

Our source of data is the Multimedia Adult English Learner Corpus (MAELC)[3] which consists of audio-visual recordings of classroom interaction in an English as a Second Language (ESL) context (Reder, Harris and Setzler 2003). The classrooms in which the recordings took place were equipped with video cameras and students were given wireless microphones on a rotational basis; the teacher also wore a microphone. Our study is a longitudinal case study consisting of

[2] We also acknowledge, with Ellis (2002), the differences between L1 and L2 acquisition and the ways in which L2 learners differ from L1 learners, including conceptual development, use of conscious problem-solving and deduction abilities, access to input and general context of learning, and L1 → L2 transfer issues.

[3] MAELC is maintained and developed under the direction of Steve Reder and Kathryn Harris at The National Labsite for Adult ESOL (known locally as the Lab School). The Lab School is supported, in part, by grant R309B6002 from the Institute for Education Science, U.S. Dept. of Education, to the National Center for the Study of Adult Learning and Literacy (NCSALL). The Lab School is a partnership between Portland State University and Portland Community College.

recordings from September 2001 through February 2005. The informant is an adult male learner of English, his native language is Spanish (he is from Mexico), and he became a US resident approx. 21 months prior to taking ESL classes. Our data, then, contrary to traditional research in second language collocational knowledge, cf. above, consist of oral classroom interaction. In our view, given the fact that FL is thought to enhance on-line language processing, verbal interaction makes for more intuitively appealing data than written language. The study has limits, however, in that we only have access to classroom interactions. The data presented here will only pertain to that context. Whatever goes on in the world outside the classroom has no bearing on this study, nor can we say anything about our informant's language use outside the classroom.

4.2. Analysis

We made queries into the MAELC database and found several students that would potentially qualify as informants in longitudinal research. We filtered the queries and ended up with approx. 40 sessions (each consisting of three hours of recordings) in which our informant is either wearing a microphone or sitting next to someone wearing a microphone. We then transcribed the recorded interactions and the resulting transcripts form the final database of our enquiry. In the process, we left out private speech (Ohta 2001) as well as interactions where the students were clearly practicing certain forms that included the negation pattern. We did some unmotivated looking in the transcripts, and we quickly realised that negated contexts would be a good place to start our analysis because negated constructions seemed to be present throughout our informant's development, and because negation is a relatively transparent form-meaning pattern. In alignment with the general research dicta of UBL, our analysis, then, is based on distributional analyses of samples of language during and across particular developmental periods (Tomasello 2000).

By operationalising FL as *recurring multi-word expressions* we attempt to make it clear that we are not working on a definition of FL based on a distinction between listed information in the lexicon and abstract rules of syntax. Rather, a MWE is a sequence of words used together for a relatively coherent communicative purpose. This definition is based on Tomasello's (2003) definition of 'construction', but whereas Tomasello, when defining construction, refers to an abstract level of language knowledge, we focus on concrete recurrences of specific words. In terms of analysing general recurrence in learner production, i.e. output frequencies, we distinguish between type and token frequency in order to capture the recurrences of both concrete expressions and more abstract patterns and constructions. In this we

reflect the UBL conceptualisation of language knowledge as a continuum of linguistic constructions, cf. section 2 above. Note also at this point that we assume that questions of frequency pertain to output as well as input. This assumption is based on the UBL principles, mentioned in section 2 above, that language usage and language knowledge as well as social interaction and individual cognition, are mutually constitutive, and that there is a link between what is conventionalised in speech community and entrenched in the kind of the individual.

5. Results

5.1. Negation pattern usage development

Figure 1 below presents an overview of the negation patterns found in the data and how their usage frequencies develop over time. Initially, we observed three distinct but related kinds of negation pattern in the data, a) recurrent target language (TL) MWE: *I don't know*; b) learner pattern: *Subj no V* (e.g., *I no remember*); c) TL pattern: *aux-neg* pattern (e.g., *I don't think so*). In 2001, negations are divided between 1/3 MWE and 2/3 learner patterns. In 2005, there is a 42-58% division between MWE and other instances of the TL pattern. In between we have varying stages of competition between the learner pattern and the TL pattern.[4]

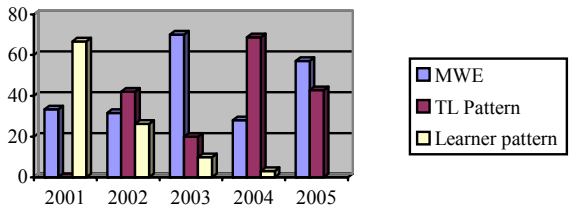

Figure 1. Negation pattern usage development.

We note three tendencies: stability and importance of the MWE, increase in TL pattern usage, and decrease (until the point of disappearance) in learner pattern usage. This gives us a rough developmental path, in terms of the language

[4] Keeping in mind Tomasello's idea of children's psycholinguistic units being different from those of adults, we note that the same might be the case for L2 learners vs. native speakers in that our informant for some time employs a distinct 'learner pattern'.

inventory, on which we tentatively pinpoint three chronologically overlapping phases:

1: TL MWE *I don't know* + Learner pattern
2: TL MWE *I don't know* + Learner pattern + TL pattern
3: TL MWE *I don't know* + TL pattern

These tentative results imply that going further into the data, we would expect to see a fairly high and stable token frequency of the MWE and an increasingly higher type frequency in the TL pattern usage across time. This would then suggest the entrenchment and whole-unit status of the MWE and the development in increasingly more abstract knowledge in terms of the TL pattern. It would also lend credence to the cohabitation in the grammar hypothesis.

5.2. Type and token frequencies

Table 1 below presents type and token frequencies as well as type-token ratios for the TL patterns and the learner pattern across time. Type frequency indicates variation in negated verbs, but it does not take into account variance in SUBJ-position; the higher the type-token ratio, the more varied and productive the usage of the pattern. In 2001, though, only one verb is negated (*know*), thus there is only one type and by implication the type-token ratio is incalculable. Although we cannot say anything definite about issues of storage, the token frequency of the TL pattern for this period is relatively high, supporting the whole-unit status of *I don't know*. This corresponds to what we expected for the relation between token frequency and MWEs. The development in type-token ratio suggests that the TL pattern is becoming increasingly more abstract in its schematic representation as productivity is expanded in terms of negated verbs. A system is emerging and the learner pattern seems to disappear.[5] Table 1 thus supports what we had expected.

[5] Please note that we are not postulating an endpoint of acquisition neither in this learner nor generally speaking. We cannot be certain that things disappear from linguistic experience. For example, one instantiation of the learner pattern in 2002 seems to be a strategic self-repair by our informant, in a situation where, using the TL pattern, he is repeatedly not understood by a fellow student. This implies that the learner pattern may be a useful interactional resource and it underlines the need for future research on how aspects of interaction, e.g., turn taking and turn constructional units in conversation (e,g, Selting 2000), interplays with the learning of the constructions of a second language.

	TL PATTERN			LEARNER PATTERN		
	TOKEN	TYPE	RATIO	TOKEN	TYPE	RATIO
2001	5	1	--	10	8	0,80
2002	14	7	0,50	5	4	0,80
2003	9	3	0,33	1	1	--
2004	31	13	0,42	1	1	--
2005	14	7	0,50	0	0	0

Table 1. Type and token frequencies for the TL pattern and the learner pattern.

Two things are puzzling, however. For instance, 2004 sees the highest amount of different negated verbs without yielding the highest type-token ratio and therefore not showing the highest level of creativity, and 2002 sees the same type-token ratio as 2005, seemingly suggesting the same level of abstractness for those two years and by implication decrease in abstractness in 2003-4. First, in the case of the level of creativity in 2004, we found that the informant not only used the MWE *I don't know* quite frequently, but other possible MWE candidates as well, namely, *I don't think so, I don't remember,* and *I don't have NP*. *I don't have NP* and *I don't remember* were not counted as MWEs in this study because they were only used in 2004. *I don't think so* comes closer to qualifying as a MWE in that it is employed by this learner from an early point in development; however, we chose not to include it as MWE because its token frequency is markedly lower than that of *I don't know*.[6] Thus, in 2004 the TL pattern was most frequently used with four recurring verbs, explaining the lower type-token ration for that period. What we cannot explain statistically is the productivity in 2002 and the idea that the informant's language inventory should be more abstract at this point than later on in development. The relatively high numbers in 2002 could be due to a number of reasons that promise interest as focal points of future research, e.g., the nature of tasks in classroom, and functional requirements in interaction. However, we may be able to explain the phenomenon by examining more closely the existing variance in the negation patterns.

5.3. Degree of abstractness in negation patterns.

Table 2 below presents the degree of abstractness in negation patterns, and how it develops across time. It shows what negation patterns are employed at what stages and with what pronouns and what tense morphology. The figure supports what was

[6] In total the type-token ratio of *I don't think so* is 0,07; for *I don't know* it is 0,44.

suggested by the type-token ratios above**,** namely**,** that the learner is working on an increasingly abstract language system. We see this in the expansion of the combinatorial possibilities. Consequently, the explanation for the high productivity / creativity in 2002 seems to be that the learner is working on a pattern which is not as abstract in its schematicity as the construction that is emerging but abstract enough to be productive. It resembles what Ellis (2002) calls a "low-scope pattern" and could be likened with what Tomasello (2003) refers to as a pivot schema. The difference between the 'pattern' and the 'construction' is thus a matter of abstractness as alluded to earlier.

	2001	2002	2003	2004	2005
1st SG	TL / LP	TL	TL / LP	TL / LP	TL
2nd SG	LP	TL / LP			TL
3rd SG	LP	LP			TL
1st PL				TL	
2nd PL					
3rd PL				TL	
Past tense				TL	TL

Table 2. Negation pattern abstractness development.

The three phases on the developmental path that we tentatively pinpointed earlier have emerged as robust and are reiterated below; this time, in terms of more elaborate constructional representation:

1: **TL MWE**: I don't know + **Learner pattern**: PRN^{sg} neg V^{prs}
2: **TL MWE**: I don't know + **Learner pattern**: PRN^{sg} neg V^{prs} + **TL pattern**: PRN^{sg}; ^{1st,2nd} aux-*do*^{prs} neg V
3: **TL MWE**: I don't know + **TL construction** PRN aux-*do*$^{prs/pst}$ neg V (S)[7]

This representation makes the increasing abstractness clear, and one notes with interest that the learner pattern is just as abstract in its representation as the TL pattern; in fact, in combinatorial terms it seems to be more productive, with the TL pattern only including 1st and 2nd person pronouns as subject. The emergence of the TL construction coincides with a more varied use of pronouns and, importantly in

[7] Arguably, the representation should not contain the 'prs/pst' notation; the 'aux' category should stand on its own, involving this possibility implicitly. We have included it to emphasise the difference in underlying representation between the learner pattern and the learner construction.

terms of abstract schematicity, it enables the learner to express past tense negation. Table 3 below is a sketchy attempt to capture this fluid development from the concrete item to the abstract system, and the relationship of the *aux-neg* construction with some other significant uses of *aux do*, showing how the negation construction seems to be used and learnt, by our informant, in a network of interrelated emergent constructions. Development is depicted horizontally, phases vertically. Repeated lexical items (intra- as well as inter-constructionally) are in bold types, MWE candidates in italics:

*I don't **know*** → *I don't think so* → I don't agree → I don't know exactly…			
↓	**Take**	need	much about
↓	remember		**have** NP
do you know (x)?			↓
yes I do / no I don't			↓
↓			↓
↓			↓
do *you **have*** x?	→ you don't **take** a shower?		→ we don't **have** NP
*You **have*** x?	**know**?		you didn't see that?
↓	go?		
↓	come to x?		
↓			
***How** do you say x*? → **how** do you write		→ **how did** you pronounce that?	
	spell		↓
			→ we **did** many things

Table 3. Overview of the inventory: The Aux-do network.

While depicting the development of the *Aux-do network*, Table 3 also reveals the need for further research in (at least) the following areas: a) usages of *can't* and *won't* in relation to the *do*-schema; b) *copula*-negation and *have*-negation in relation to *do*-negation; c) past tense usage of the *do-schema* in relation to other usages of past tense; d) development of the *PRN* in the *do*-schema in relation to other pronoun usages; e) the role of other MWEs in other constructions; and f) other usages of the negated verbs.

6. Conclusions

The analyses presented above indicate that from the MWE *I don't know* emerge increasingly abstract patterns which sanction systematic usage expansion of the negation pattern to include other lexical items, i.e. varied usage of PRNs and main verbs, and, perhaps more importantly in terms of level of abstractness, past tense expression. The system that is seen to emerge in this fashion is the gradual abstraction of regularities that link expressions as constructions. Second language learning may thus be seen as exemplar-based (Ellis 2002).

Furthermore, our data support two key UBL principles. The first relates to notions of frequency; productivity of a given constructional schema seems to be dependent on type frequency, and entrenchment of a given MWE seems to be dependent on token frequency. In our data, the instantiations of the *aux-neg* schema got more varied across time, whereas the expression *I don't know* was retained throughout development as a MWE. This supports our assumption that questions of frequency pertain to output as well as input. The second principle is that language knowledge, acquisition, and storage may be dually represented, cf. the rule/list fallacy. In our data *I don't know* as a MWE is stored as a whole that, in turn, becomes sanctioned by the very schema whose emergence it initiates.

Finally, our study does not tally well with the idea of MWEs as syntactic freezes, as a MWE in our conceptualisation is not seen as a representation of underlying syntactic rules and not necessarily schematised by underlying abstract language knowledge. Furthermore, the elements of the MWE *I don't know* cannot be thought of as frozen in this pattern as they are employed elsewhere by the learner; initially, though, the negation element is not. This does not mean, however, that the ability to negate linguistic material is frozen in this one initial instantiation of the *aux-neg* construction; remember that the learner is perfectly capable of negating in general; initially, he merely uses another pattern to do so productively, namely the learner pattern *subj no V*. Nothing is frozen here; rather, the MWE forms the backbone of schematic development, and across time it becomes a reflection of the emergent abstract language knowledge.

References

Achard, Michel. 2006. Usage-based semantics: Meaning and distribution of three French 'breaking' verbs. Paper presented at Collocations and Idioms: The First Nordic Conference on Syntactic Freezes, Joensuu 19-20/5-06.

Cadierno, Teresa. 2004. Expressing motion events in a second language: A cognitive typological approach. *Cognitive linguistics, second language acquisition and foreign language pedagogy*, ed. by Michel Achard & Susanne Niemeier, 13-49. Berlin: Mouton de Gruyter.

Cadierno, Teresa and Lucas Ruiz. 2006. Motion events in Spanish L2 acquisition. *Annual Review of Cognitive Linguistics* 4. 183–216.

Ellis, Nick C. 2002. Frequency effects in language processing – a review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24. 143-188 .

Gitsaki, Christina. 1999. *Second language lexical acquisition. A Study of the Development of Collocational Knowledge*. San Francisco: International Scholars Publications.

Hopper, Paul J. 1998. Emergent grammar. *The new psychology of language, volume 1*, ed. by Michael Tomasello, 155-175. Mahwah, NJ.: Lawrence Erlbaum

Langacker, Ronald W. 1987. *Cognitive grammar, vol. 1. Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Nesselhauf, Nadja. 2005. *Collocations in a learner corpus*. Amsterdam: John Benjamins.

Ohta, Amy S. 2001. *Second language acquisition processes in the classroom*. Mahwah, NJ.: Lawrence Erlbaum

Reder, Stephen, Kathryn A. Harris, and Kristen Setzler. 2003. A multimedia adult learner corpus. *TESOL Quarterly* 37, 3. 546-557.

Schmitt, Norbert. (ed.) 2004. *Formulaic sequences*. Amsterdam: John Benjamins.

Selting, Margret. 2000. The construction of units in conversational talk. *Language in Society* 29. 477-517.

Tomasello, Michael. 2000. First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics* 11-1/2. 61-82.

Tomasello, Michael 2003. *Constructing a Language*. Cambridge: Cambridge University Press.

Tummers, Jose, Kris Heylen, and Dirk Geeraerts. 2005. Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1-2. 225-261.

Van Lancker-Sidtis, Diana and Gail Rallon. 2004. Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language and Communication* 24. 207-240.

Waara, Renee. 2004. Construal, convention, and constructions in L2 Speech. *Cognitive linguistics, second language acquisition and foreign language pedagogy*, ed. by Michel Achard and Susanne Niemeier, 51-75. Berlin: Mouton de Gruyter.

Wray, Alison. 2002. *Formulaic language and the lexicon*. New York, NY: Cambridge University Press.

Søren Wind Eskildsen
Institute of Language and Communication
University of Southern Denmark
Campusvej 55
5230 Odense M
Denmark
swe@language.sdu.dk
http://www.humaniora.sdu.dk/nywebX/inc/show.php?full=1563

Teresa Cadierno
Institute of Language and Communication
University of Southern Denmark
Campusvej 55
5230 Odense M
Denmark
cadierno@language.sdu.dk
http://www.humaniora.sdu.dk/nywebX/inc/show.php?full=88