

Volume 27:2014

**sky**

## **SKY Journal of Linguistics**

Editors:

Markus Hamunen, Tiina Keisanen, Hanna Lantto,  
Lotta Lehti, Saija Merke

*SKY Journal of Linguistics* is published by the Linguistic Association of Finland (one issue per year).

### Notes for Contributors

**Policy:** *SKY Journal of Linguistics* welcomes unpublished original works from authors of all nationalities and theoretical persuasions. Every manuscript is reviewed by at least two anonymous referees. In addition to full-length articles, the journal also accepts short (3–9 pages) ‘squibs’ as well as book reviews.

**Language of Publication:** Contributions should be written in English, French, or German. If the article is not written in the native language of the author, the language should be checked by a qualified native speaker.

**Style Sheet:** A detailed style sheet is available from the editors, as well as via WWW at <http://www.linguistics.fi/skystyle.shtml>.

**Abstracts:** Abstracts of the published papers are included in *Linguistics Abstracts* and *Cambridge Scientific Abstracts*. The published papers are included in *EBSCO Communication & Mass Media Complete*. *SKY JoL* is also indexed in the *MLA Bibliography*.

#### Editors’ Addresses (2014):

Markus Hamunen, School of Language, Translation and Literary Studies, Kalevantie 4, FI-33014, University of Tampere, Finland

Tiina Keisanen, Faculty of Humanities, P.O. Box 1000, FI-90014 University of Oulu, Finland

Hanna Lantto, Department of Modern Languages, P.O. Box 24, Unioninkatu, FI-00014 University of Helsinki, Finland

Lotta Lehti, French Department, Koskenniemenkatu 4, FI-20014 University of Turku, Finland

Saija Merke, Department of Finnish, Finno-Ugrian and Scandinavian Studies, P.O. Box 3, Fabianinkatu 33, FI-00014 University of Helsinki, Finland

**Editors’ E-mail:** [sky-journal\(at\)helsinki\(dot\)fi](mailto:sky-journal(at)helsinki(dot)fi)

#### Publisher:

The Linguistic Association of Finland

c/o General Linguistics

P.O. Box 24 (Unioninkatu 40)

FI-00014 University of Helsinki

Finland

<http://www.linguistics.fi>, <http://www.linguistics.fi/skyjol.shtml>

The Linguistic Association of Finland was founded in 1977 to promote linguistic research in Finland by offering a forum for the discussion and dissemination of research in linguistics, both in Finland and abroad. Membership is open to anyone interested in linguistics. The membership fee in 2013 was EUR 25 (EUR 15 for students and unemployed members). Members receive *SKY Journal of Linguistics* gratis.

Cover design: Timo Hämäläinen 1999





# SKY Journal of Linguistics

## 27

Suomen kielitieteellisen yhdistyksen aikakauskirja  
Tidskrift för den Språkvetenskapliga föreningen i Finland  
Journal of the Linguistic Association of Finland

### Editors:

Markus Hamunen  
Lotta Lehti

Tiina Keisanen  
Saija Merke

Hanna Lantto

### Copy editor:

Jouni Harjumäki

### Advisory editorial board:

Werner Abraham  
*University of Vienna*

Kimmo Granqvist  
*University of Helsinki*

Auli Hakulinen  
*University of Helsinki*

Martin Haspelmath  
*Max Planck Institute of  
Evolutionary Anthropology*

Marja-Liisa Helasvuo  
*University of Turku*

Anders Holmberg  
*Newcastle University*

Tuomas Huomo  
*University of Turku*

Jaakko Hämeen-Anttila  
*University of Helsinki*

Juhani Härmä  
*University of Helsinki*

Fred Karlsson  
*University of Helsinki*

Seppo Kittilä  
*University of Helsinki*

Meri Larjavaara  
*Åbo Akademi University*

Jaakko Leino  
*University of Helsinki*

Marja Leinonen  
*University of Helsinki*

Matti Miestamo  
*University of Helsinki*

Jussi Niemi  
*University of Eastern  
Finland*

Urpo Nikanne  
*Åbo Akademi University*

Martti Nyman  
*University of Helsinki*

Krista Ojutkangas  
*University of Turku*

Mirja Saari  
*University of Helsinki*

Helena Sulkala  
*University of Oulu*

Kari Suomi  
*University of Oulu*

Ulla Tuomarla  
*University of Helsinki*

Maria Vilkuna  
*Institute for the Languages  
of Finland*

Jussi Ylikoski  
*UiT The Arctic University  
of Norway*

Jan-Ola Östman  
*University of Helsinki*

2014

ISSN-L: 1456-8438  
ISSN: 1456-8438 (Print)  
ISSN: 1796-279X (Online)

Hansaprint, Turku 2014

## Contents

External Reviewers of <i>SKY JoL</i> 27 (2014).....	5
<i>Thomas Groß</i> Some Observations on the Hebrew Desiderative Construction – A Dependency-Based Account in Terms of Catenae .....	7
<i>Wojciech Lewandowski</i> Deictic Verbs: Typology, Thinking for Speaking and SLA.....	43
<i>Ulla Vanhatalo, Heli Tissari &amp; Anna Idström</i> Revisiting the Universality of Natural Semantic Metalanguage: A View through Finnish.....	67
<i>Mari Wiklund</i> La transmission des effets stylistiques des phrases sans verbe fini dans les traductions finnoises – L'exemple des pièces de théâtre de Jean-Paul Sartre.....	95
<i>Peter Wikström</i> #srynotfunny: Communicative Functions of Hashtags on Twitter .....	127
<b>Squibs:</b>	
<i>María Luisa Carrió-Pastor</i> Cross-cultural Variation in the Use of Modal Verbs in Academic English.....	153
<b>Book reviews:</b>	
Maurer, Philippe (2009) <i>Principense (Lung'le). Grammar, Texts, and Vocabulary of the Afro-Portuguese Creole of the Island of Príncipe, Gulf of Guinea</i> . Reviewed by Angela Bartens.....	167
Susana S. Fernández & Johan Falk (eds.) (2014) <i>Temas de gramática española para estudiantes universitarios, Una aproximación cognitiva y funcional</i> . Reviewed by Ilpo Kempas.....	173
Harvey, Kevin (2013) <i>Investigating Adolescent Health Communication. A Corpus Linguistics Approach. Research in Corpus and Discourse</i> . Reviewed by Pirjo Salomaa.....	179



## External Reviewers of *SKY JoL* 27 (2014)

The following scholars, among a few others wishing to remain anonymous, have acted as external reviewers for *SKY Journal of Linguistics* in 2014:

Adi Yasran Abdul Aziz (Universiti Putra Malaysia), L.O. Adewole (Obafemi Awolowo University), Lotta Aunio (University of Helsinki), Pauli Brattico, Mark Dingemanse (MPI for Psycholinguistics), Małgorzata Fabiszak (Adam Mickiewicz University, Poznań), Noam Faust (Hebrew University of Jerusalem), Rania Habib (Syracuse University), Jaakko Hämeen-Anttila (University of Helsinki), Thomas Hanneforth (Universität Potsdam), Larry Hyman (University of California at Berkeley), András Imrényi (Jagiellonian University Kraków), Anni Jääskeläinen (University of Helsinki), Olaf Koenemann (Radboud University), Aino Koivisto (University of Helsinki), Marcin Kuczok (University of Silesia), Dennis Kurzon (University of Haifa Jerusalem), Veronika Laippala (University of Turku), François Lareau (Université de Montréal), Meri Larjavaara (Åbo Akademi), Jaakko Leino (University of Helsinki), Stephen Levinsohn (SIL International), Pekka Lintunen (University of Turku), Marit Lobben (University of Oslo), Geda Paulsen (Institute of the Estonian Language), Andrea Pešková (Universität Hamburg), Pekka Posio (Universität zu Köln), Alain Rabatel (Université de Lyon 1), Susanna Shore (University of Helsinki), Dick Smakman (University of Leiden), Karen Sullivan (University of Queensland), A.M. Tessier (University of Alberta), Heli Tissari (University of Eastern Finland), Catherine Travis (Australian National University), Mari Wiklund (University of Helsinki), Maria Vilkuna (University of Helsinki), Michele Zappavigna (University of New South Wales (UNSW), Australia), Dr Zawada (University of South Africa (UNISA))



**Thomas Groß**

## **Some Observations on the Hebrew Desiderative Construction – A Dependency-Based Account in Terms of Catenae<sup>1</sup>**

### **Abstract**

The Modern Hebrew (MH) desiderative construction must obey four conditions: 1. A subordinate clause headed by the clitic *še*= ‘that’ must be present. 2. The verb in the subordinate clause must be marked with future tense. 3. The grammatical properties genus, number, and person tend to be specified, i.e. if the future tense affix is underspecified, material tends to appear that aids specification, if contextual recovery is unavailable. 4. The units of form that make up the constructional meaning of the desiderative must qualify as a catena. A catena is a dependency-based unit of form, the parts of which are immediately continuous in the vertical dimension. The description of the individual parts of the desiderative must address trans-, pre-, and suffixes, and cliticization. Catena-based morphology is representational, monostratal, dependency-, construction-, and piece-based.

### **1. Purpose, means and claims**

The main purpose of this paper is to analyze the Hebrew desiderative construction. This construction is linguistically interesting and challenging for a number of reasons. 1. It is a periphrastic construction, with fairly transparent compositionality. 2. It is transclausal, i.e. some parts of the construction reside in the main clause, and others in the subordinated clause. The complementizer is also part of the construction. 3. The construction consists of more than one word, but it does not qualify as a constituent. Rather the construction cuts into words. 4. Two theoretically

---

<sup>1</sup> I want to thank Outi Bat-El (Tel Aviv University) and three anonymous reviewers for their help and advice. Statements made in this paper do not necessarily reflect their positions. Any mistakes remain my responsibility.

challenging phenomena are in play: transfixes, and clitics. These aspects are illustrated with the next example:<sup>2</sup>

- (1) *hu rotse še= ʔani ʔe-ftor xid-ot.*  
 3SG.M want.SG COMP- 1SG FUT.1SG-solve.IRR riddle-PL  
 ‘He wants me to solve riddles.’

The boldface letters are the surface-based units that make up the desiderative construction. The boldface consonants, *r...ʔ...*, in the second word mark the construction root, a verb marked for volition.<sup>3</sup> The vowels *...o...e*, a transfix, are not part of the construction, because, for obvious reasons, the verb of volition may appear in any tense, mood, person, genus, number, or any combination thereof.<sup>4</sup> The distinction into transfixes (patterns) and radicals (roots), and how this distinction is represented in the dependency grammatical tree representations is addressed in section 3.1.<sup>5</sup>

The verb in the subordinated clause must be marked with the future tense (cf. section 3.2). In Hebrew, future tense is prefixed, and these prefixes differ in how “cumulative” they are. The prefix *ʔe-* in (1) expresses tense (future tense), person (first person), and number (singular). It does not express genus, though. The lexical verb in the subordinated clause, indicated by the consonants *...ft...r*, sits in a slot opened up by the construction at the bottom. But it does not contribute to the grammatical meaning expressed by the construction, rather it provides content. A sentence almost identical to example (1) is analyzed in section 4 as example (25).

The item *še=* ‘that’ is a clitic, and it functions as a complementizer. It is attached to the first word of the subordinated clause, which here is *ʔani*

<sup>2</sup> I adopt a relative standard transcription of Modern Hebrew, with *h* and *ʔ*, which are often not pronounced (where *ʔ* is a merger of the historical *ʔ* and *ʕ*), and the first person *ʔe-*, which is increasingly being replaced with the 3rd person prefix *ji-*. *ʔ* stands for /ʔs/.

<sup>3</sup> The symbol ‘...’ is used for material that interrupts a catena in the horizontal dimension. See section 2.1.

<sup>4</sup> One anonymous reviewer remarks that some material appearing above the radical *can* influence whether a desiderative construction can be governed by the volitive radical. *HITPA'EL* forms of volitive verbs can result in passivized verbs, and since volitives as such cannot be passivized, desiderative constructions fail to be grammatical in these contexts. This, however, concerns valence and voice, but not tense, mood, genus, person, or number.

<sup>5</sup> While the term is “root-and-pattern” morphology, I will use “radical” instead of “root”, because the latter carries a different meaning in this theory, namely the topmost node in a tree structure or construction. Sometimes I use “root” because it is preferable.

‘I’. This pronoun is not part of the construction, rather it is used here to show that material not germane to the construction can appear between parts of the construction. The pronoun is not necessary, because the prefix *ʔe-* already expresses the respective grammatical meaning. It is important to note that the clitic *še=* ‘that’ and the pronoun *ʔani* ‘I’ form one prosodic word, but these two items do not entertain a syntactic relationship with one another. Rather the clitic dominates the verb in the subordinated clause, which dominates the pronoun. Cliticization is addressed in section 3.3.

The analysis to be proposed below is surface-, dependency-, and catena-based, and entirely representational. Operating exclusively on the surface, without acknowledging hidden levels of representation, leads to a piece-based theory of morphosyntax. This means that the current account acknowledges units of form smaller than the word. Such a unit is called “morph”. Morphs constitute individual nodes of morph catenae. A catena is a scalable unit of form that comprises any immediately connected surface units in the dominance dimension. The required notions and terms are introduced in section 2.

The principal claims made here are:

1. The units of form that make up the Hebrew desiderative construction qualify as a catena, rather than as a word or a constituent.
2. The Hebrew desiderative is a construction that cuts into words.

A catena-based analysis of the Hebrew desiderative is attainable, once the groundwork for analyses within dependency morphology is laid. The analyses of the parts of the desiderative construction (section 3) are therefore integral to the entire argument. A secondary purpose of the paper is to develop a general understanding of the crucial notions of catena-based dependency morphology, and to illustrate how morphological relationships are represented.

The paper proceeds as follows: Section 2 provides general information on dependency grammar, and the notion of the catena. It also gives a historical account of dependency morphology. It then lays out morphological notions based on the catena. Finally it formulates several reservations against other approaches to morphology. Section 3 introduces the principal players involved in the Hebrew desiderative construction: transfixes and radicals (3.1), the future tense prefix system (3.2), and

cliticization (3.3). Section 4 then combines these phenomena in order to analyze the desiderative construction. Section 5 summarizes the paper.

## 2. Theoretical background

The proposal to be made here operates within dependency grammar. This framework originates with Lucien Tesnière (1959), and has produced a considerable body of literature. The following list is by no means complete, but it reflects the historical development of the field: Hays (1964); Robinson (1970); Kunze (1975); Matthews (1981); Sgall, Hajičová & Panevová (1986); Mel'čuk (1988); Schubert (1988); Starosta (1988); Lobin (1993); Pickering & Barry (1993); Engel (1994); Jung (1995); Heringer (1996); Groß (1999); Eroms (1985, 2000); Kahane (2000); Tarvainen (2000).<sup>6</sup> Richard Hudson's *Word Grammar* (1984, 1990, 2007, 2010) has contributed significantly to making dependency grammatical concepts known. Many detailed introductions and discussions of specific issues pertinent to dependency grammar can be found in Ágel, Eichinger, Eroms et al. (2003, 2006). In recent decades, computer linguistics, too, has increasingly looked toward dependency (Nivre 2006).

The fundamental and overarching properties of dependency grammars are that they are word-based, and that they regard the dependency relation between words as basic. While phrases or constituents are at times involved in the analysis, they are not considered basic units of such an analysis. Apart from that, dependency grammars come in several flavors: derivational or representational, mono- or multistratal, construction- or rule-based. The account here is representational, monostratal, and construction-based.

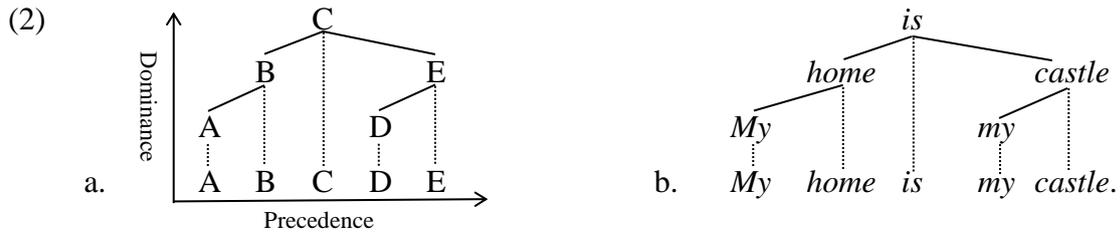
This section provides a brief overview over catena-based dependency grammar (section 2.1), and also brief history of dependency morphology (section 2.2). In section 2.3, the extension of catenae into morphology is demonstrated. Section 2.4 briefly remarks on competing theories, in particular word/lexeme/paradigm-based accounts of morphology.

---

<sup>6</sup> Dependency-based "Meaning-Text Theory (MTT)", founded by Igor Mel'čuk, has attracted a large number of linguists.

## 2.1 Catena-based dependency grammar

The current account is closely associated with the concept of the *catena*. Consider first the next representations:



Representation (2a) is a virtual dependency tree. It is virtual because, instead of actual words, capitals appear as nodes. Two dimensions are distinguished: the horizontal dimension (x-axis) is called *precedence* because in this dimension the linear order of the nodes appearing in the tree structure is established. The vertical, dotted edges are called *projection edges*, and they show the order in which the individual nodes project to their position in the example shown at the bottom. The vertical dimension (y-axis) is called *dominance*: in this dimension the dominance relationships between the nodes are represented by angled, solid *dependency edges*. This type of tree representation is most closely associated with the work of Hays (1964). Tree representations in other dependency theories may look quite different, as they may concentrate only on one dimension. Tree (2b) is a “real-life” example the syntactic dependency structure of which is equivalent with that shown in (2a). For the most part, the individual dependencies are based on assumptions similar to constituency structure. The words *My home* constitute a noun phrase because the entire expression behaves like a noun. In dependency grammars the article thus depends on the noun.<sup>7</sup> The tree root is the verb *is*, rather than an exocentric node *S*.<sup>8</sup>

The term *string* is a unit with respect to the precedence dimension. Every node, and every combination of nodes that is continuous, i.e. uninterrupted, in this dimension is called a string. In (2a), the following units qualify as strings:

<sup>7</sup> Apart from Word Grammar, most dependency grammars reject the idea of a DP.

<sup>8</sup> This property may be the overarching distinction between dependency and constituency grammars. For a discussion of finite vs. non-finite VPs, see Osborne, Putnam & Groß (2011: 323).

- (3) A, B, C, D, E, AB, BC, CD, DE, ABC, BCD, CDE, ABCD, BCDE, and ABCDE

Insert the individual words in (2b) for the respective capitals in (3), and one gets the set of strings valid for (2b).

The term *catena* is a unit with respect to the dominance dimension. Similar to the concept of string, every node, and every combination of nodes that is continuous, i.e. uninterrupted, in the dominance dimension is called a catena. In (2a), the following units qualify as catenae:

- (4) A, B, C, D, E, AB, BC, C...E, DE, ABC, BC...E, CDE, ABC...E, BCDE, and ABCDE

A comparison of (3) and (4) reveals that the node combinations CD, BCD, and ABCD only qualify as units in the precedence dimension, i.e. they are strings. The node combinations C...E, BC...E, and ABC...E, however, do NOT qualify as strings (because node D interrupts the continuity in the precedence dimension), yet these node combinations do qualify as catenae because they are uninterrupted in the dominance dimension. On the other hand, the string node combinations CD, BCD, and ABCD do NOT qualify as catenae because the node E interrupts the continuity in the dominance dimension.

Even though the proposal below uses two-dimensional tree representations, the term *string* and *catena* allow one to talk about relationships in different dimensions in isolation. A dependency grammar is particularly suited to visualize catenae. Recent research has established that the catena is centrally involved in a number of grammatical phenomena that have, over the decades, challenged theories of grammar, in particular constituency-based theories. Based on the precursor to the catena, O’Grady’s (1998) “chain”, Osborne (2005) introduces the notion to dependency grammar. Groß and Osborne (2009) use this notion in order to explain displacement and related phenomena, among them *w(h)*-fronting, topicalization, scrambling, extraposition, inversion, shifting, free relatives, and pied-piping. Osborne, Putnam and Groß (2012) introduced the label “catena” in order to avoid confusion with other uses of the term “chain” in linguistics. This paper discusses a number of pertinent properties of the catena, and it compares the concept to the constituent. Most importantly, the paper establishes that the catena plays a central role in analyzing idiom formation, ellipsis (answer fragments, gapping, stripping, VP ellipsis, pseudo-gapping, sluicing, comparative deletion), and predicate structure. Osborne, Putnam and Groß (2011) attempt a reevaluation of developments

within Minimalism in light of the concept of the catena, arguing that the latest Minimalist versions converge on core concepts of dependency grammars, though they also point to the limits of such developments. Osborne and Groß (2012a) argue that “constructions”, as posited in Construction Grammar, can be recovered as catenae. Further, Osborne and Groß (2012b) argue that antecedent-containment can be parsimoniously explained when utilizing the catena concept.

## 2.2 Dependency morphology

While research on dependency-based syntax can draw on an extensive body of literature, contributions on dependency-based morphology are difficult to find. The earliest attempt at describing morphological structure with dependencies can be found in Heringer (1970: 96).<sup>9</sup> He has been using dependency-based morphological trees consistently, but sparingly (1973: 283–294, 1996: 117–118). The name “dependency morphology” was originally proposed by John Anderson (1980) in a dependency-based analysis of the Basque verb. The general dearth of dependency-based morphological analyses is lamented in Harnisch (2003) and Maxwell (2003). But what might the causes of this dearth be?

One reason has to do with the rigor of analysis. The unifying aspect of John Anderson’s, Heringer’s, and others’ analyses is the assumption that affixes depend on lexical material. This assumption, however, conflicts with a significant body of knowledge accumulated since Joseph Greenberg. Bybee (1985) makes the compelling point that there is a hierarchy that orders the appearance of derivational and inflectional affixes on the verb. This hierarchy has semantic significance, and hence the assumption should be that the affixes dominate the lexical material. A similar observation accounts for nouns marked for number and/or case. The dependency-based attempts at morphology do not take the basic insight associated with Bybee’s hierarchy into account. Assuming analyses that have affixes dominating lexical material conflicts, however, with a tenet held to be central by many dependency grammarians, namely the concept of valency. One wishes to see lexical material as the root (node) in order to maintain a

---

<sup>9</sup> This early date is astonishing given the fact that hierarchical word structure within constituency-based theories of morphology is first proposed by Williams (1981).

valency-oriented description.<sup>10</sup> As long as valency is seen as central, the assumption that non-lexical material is somehow subsumed by lexical material seems logical.

Another factor that surely has contributed to scotching the development of dependency-based morphologies is the upswing of word/lexeme/paradigm-based morphology (Robins 1959; Matthews 1972; Aronoff 1976; Spencer 1991; Anderson 1992; Stump 2001; Booij 2010; Stump & Finkel 2013; and many others) making the case against piece-based morphology. Proposals such as those by Stephen Anderson (1992) and Stump (2001) go so far as to reject both the necessity and the possibility of segmenting words into individual morphemes. Of course granting credence to such a stance obviates any approach that sees internal word structure as similar to sentence structure.

Yet another reason has to do with a core concept of dependency grammar. Dependency grammars seem, by their very nature, to be word grammars. If words are seen as the basic units of syntax, then their further analysis into component parts is deemed inappropriate, at least as long as it concerns syntax. MTT and Word Grammar are a case in point: while they account for morphological structure in detail, their morphologies always see the word as such as the domain within which these matters play out. While Mel'čuk (1988: 107, 2003: 193) acknowledges morphological dependencies, he also delimits them from dependency structure proper. And the networks assumed in Creider and Hudson (1999), and Hudson (2003: 514, 518; 2007: 63–116) purport to illustrate the interaction and realization of features, but these are encapsulated within the word itself. As Hudson (2010: 132) notes “[m]orphology...describes changes within a word”. This stance forecloses the possibility of viewing the interplay between syntax and morphology as a continuum, and has thus helped to reinforce the view of morphology encapsulated from syntax.

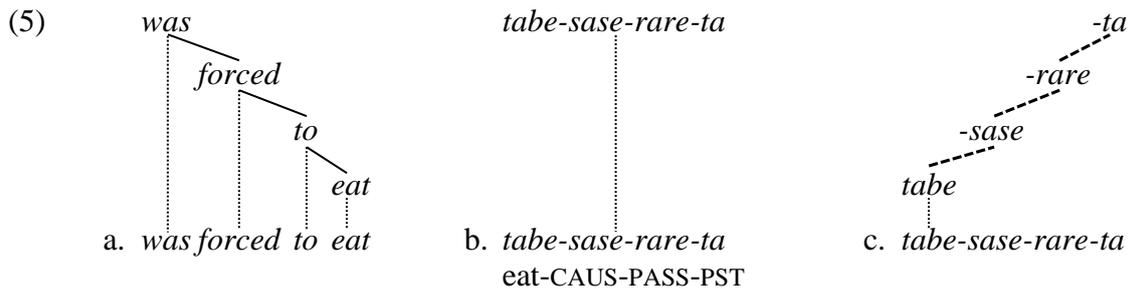
To summarize, dependency morphology has floundered (until now) due to at least four considerations: analytic errors, overly strict adherence to valency theory, the influence of word/lexeme/paradigm-based morphology, and the emphasis on the word.

---

<sup>10</sup> Especially in the European tradition, dependency grammar is seen as a supporting theory to valency theory. Ágel and Fischer (2010), for instance, devote 14 pages to valency theory but only 8 pages to dependency-based hierarchical organization of linguistic units.

### 2.3 Catenae in morphology and morphosyntax

The smallest catena consists of one node. In syntax, nodes are words. But “word” is a language-specific unit. At times, a purely word-based analysis is unsatisfactory. Consider the next examples, one from English, and its Japanese equivalent:



The English example (5a) contains four words, but the Japanese example (5b) contains only one word. The grammatical meanings causative and passive are expressed in Japanese by suffixes. In order to better compare, in particular, the dominance structure of both expressions, the Japanese word needs to be broken down into its meaning-bearing parts. (5c) shows the word-internal tree structure of the Japanese example.

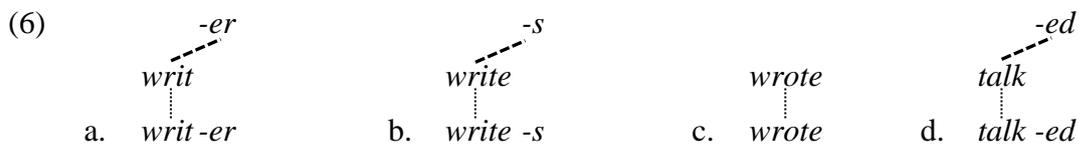
A node at the morphological or morphosyntactic level is called a *morph*. A morph is a unit of form, **not** a unit of meaning. A morph need not express exactly one unit of meaning, but rather it may – and often does – express complex meaning. No attempt at **always** matching exactly one unit of form to exactly one unit of meaning is made here.

If an expression can be reduced so that the remainder of the expression expresses a part of the entire meaning of the expression, and if the remainder cannot be reduced any further, then this remainder is a morph. E.g. German [machst] ‘(2SG) do’ can be reduced to yield [st], which expresses [2SG].<sup>11</sup> [st] cannot be reduced further without compromising the meaning of the entire expression. Hence [st] qualifies as a morph. If this morph is reduced from the entire expression, [mach] remains, which expresses part of the entire expression, and which cannot be reduced any further either. Hence, [mach] is also one morph. It is evident that this

<sup>11</sup> It is irrelevant whether one wishes to ascribe to [st] the meanings of tense or of mood. For one thing, it would only enhance the meaning expressed by [st], and since morphs are allowed to express complex meaning, no counter-argument is present. However, it is also not evident whether such an ascription would be accurate.

approach must proceed carefully, and in a conscientious manner. A second example illustrates this attitude. English *went* is considered as one morph because reduction is not possible, even though it expresses more than one meaning. An analysis of *went* as ‘go’ + PST is viewed here as a semantic analysis, not an analysis of the structure.

Immediate dominance relationships between morphs are justified by distribution. In this respect, it is irrelevant whether an affix is inflectional or derivational. This is demonstrated below with English examples:



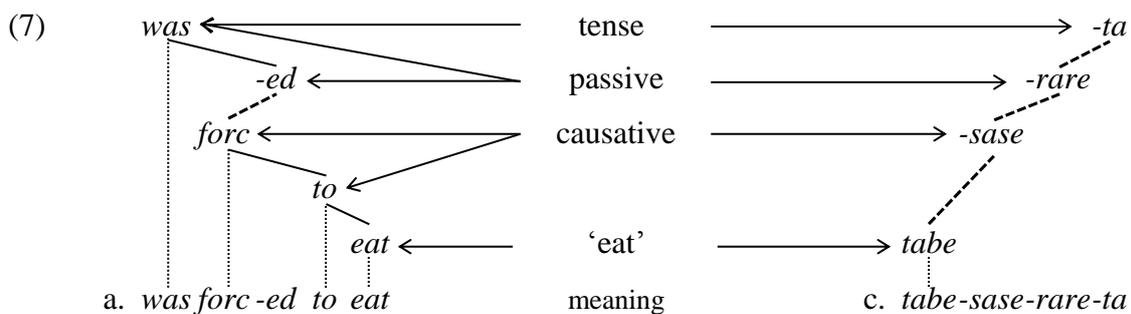
The expression *writ(e)* in (6a) is viewed as dominated by the derivational suffix *-er* because even though it is a verb, the entire expression, i.e. *writer*, distributes like a noun. For this distribution the suffix is responsible: *a writer, the writer, famous writers*, rather than a verb: *\*quickly writer, \*would writer*, etc. In (6b), the non-past third person singular suffix *-s* causes the distribution to vary from that of *write* alone: *he/she writes* vs. *\*he/she can/has writes*. The verb *write* is irregular, since its past tense form is *wrote*. The expression *wrote* is viewed as **one** morphological node even though it is semantically complex, i.e. *write* + PST, and appears in the same paradigm as *talk-ed* in (6d). In (6d) the suffix *-ed* is viewed as dominating *talk*, because the suffix specifies the distribution of the entire expression *talked*, regardless of whether one wishes to see the inflectional past tense suffix, or the derivational past participle suffix *-ed*. This kind of approach is, of course, not limited to verbs, but can apply to all kinds of lexical and affix material.

It should be kept in mind, though, that this approach rests on the crucial notion that catenae are expressions of meaning. It does not matter whether a catena is simplex or complex, or whether the meaning expressed is simplex or complex. Nodes, be they syntactic or morphological, do not always produce compositional meaning. In syntax, expressions of non-compositional meaning are called idioms, if the meanings are lexical, and periphrastic constructions, if the meanings are grammatical. A catena-based analysis of idioms is proposed by Osborne, Putnam and Groß (2012). Osborne and Groß (2012a) argue that periphrasis should be analyzed as catenae. Finally, Groß and Osborne (2013) argue that periphrastic

constructions reach into words, and that they should be, as a result, analyzed as morph catenae.

The underlying theory here has been proposed by Groß (2010, 2011a, 2011b, 2014). Groß (2010) argues that internal word structure is not fundamentally different from syntactic structure. Groß (2011a) outlines the foundations of dependency relations within and across words. Groß (2011b, 2014) addresses the topic of cliticization. Groß and Osborne (2013) illustrate that constructions most often cut into words, i.e. a (periphrastic) construction does not qualify as a word combination, nor a constituent, but rather it qualifies as a catena, the individual nodes of which are parts of words. Most importantly, they argue that the catena is a scalable unit of form that can be applied to phenomena ranging from syntax to morphosyntax to morphology proper.

Two types of dependencies are distinguished: intra-, and inter-word dependencies.<sup>12</sup> Intra-word dependencies exist between morphs belonging to the same word, while inter-word dependencies exist between morphs of two (or more) different words. Tree (5c) is an example showing only intra-word morphological dependencies. Below, example (5a) is considered again as (7a). It illustrates an inter-word dependency in English. The arrows are symbols of meaning ascription, the labels of the meanings are shown in the center of the representation.



The only difference between (5a), and (7a) is that *forced* is now shown as two nodes.<sup>13</sup> The morph catena *forc-ed* is an instance of an intra-word dependency because this expression qualifies as one word. The morph

<sup>12</sup> The distinction and the respective definitions for the two relationships are not given here. Groß (2011a) provides a detailed account.

<sup>13</sup> It is irrelevant here whether *forced* would better be rendered as *force-d* because it would not impact the argument made here. The Latin alphabet, instead of the phonetic alphabet, is used here for convenience.

catena *was...-ed* is an instance of an inter-word dependency because these two morphs do not qualify as one word, but rather belong to different words.

The top node *was* is involved in the expression of tense and voice. The morph catena *was...-ed* is the expression of a periphrastic construction, namely the English passive. The morph catena *forc...to* is the expression of causative, the *to* being necessary in order to dominate another verb.

If represented in this fashion, a comparison of the (grammatical) meaning structures of the two vastly different languages appears much more promising. In this light, periphrastic constructions, which are a challenging topic for any theory of grammar, are morph catenae across two (or more) words.

## 2.4 Remarks on other approaches to morphology

It is conventional to put forth at least some reservations against theoretical notions one opts not to follow, even though it is hoped that the explanations in section 2.3, and in sections 3 and 4 will actively demonstrate that a piece-based account utilizing the catena can, in fact, deal with phenomena that are difficult to address in any framework. The rejection of the word/lexeme/paradigm-based approaches to morphology must, for the sake of brevity, rest on two issues: the difficulties of capturing the expression of non-compositional morphology on the surface, and bracketing paradoxes.<sup>14</sup>

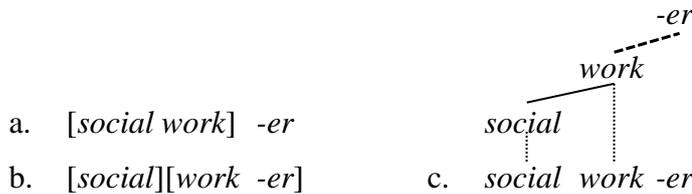
According to Matthews (1972, 1991: 201) Priscianic, or parasitic, formation occurs when a form appears only as an attachment site for other material but fails to express the meaning(s) that it would have in isolation, or with yet other material. The prime example in the literature is the Latin future participle. Here an example from Aronoff (1992: 6):

---

<sup>14</sup> In addition, several morphological phenomena, such as transfixes, circumfixes, infixes, suprafixes, reduplication, etc. have been put forward against piece-based analyses. Transfixes are discussed in section 3.1. I do not discuss the remaining phenomena here because I have done so already elsewhere (Groß 2011a) thereby showing that these issues are not insurmountable in a piece-based approach.



(9)



The expression *social worker* refers to a person who engages in social work, i.e. we usually understand the work of the person to be of a social nature, rather than the person. (9a) would thus represent the correct bracketing structure for this meaning attribution. It is seen as problematic, however, that in order to do so, one must acknowledge that the unit [social work] cuts into the word *work-er*. This assumption, namely, conflicts with the Lexical Integrity Hypothesis, which states that morphology is inaccessible to syntax.<sup>17</sup> Hence, structure (9b) is widely regarded as correct. The paradoxical nature of the conflict between (9a) and (9b) increases when the number of units of form is compared.

- (10) a. social, social work, social worker = (9a)  
 b. social, work, social worker = (9b)  
 c. social, social work, work, worker, -er, social worker = (9c)

The catena-based representation (9c) is preferable over the constituent-based analyses (9a or 9b) because it not only describes the correct semantic relationships, but it also identifies all units of form (10c). That is the result of the catena being a more inclusive unit of form than the constituent.<sup>18</sup> The blind spot of constituent structure is the inability to accurately single out units of form in the vertical, i.e. dominance, dimension.<sup>19</sup> If morphology only analyzes phenomena in the horizontal, i.e. precedence, dimension, and utilizes only constituent structure whenever structuring is desired, then these problems become predictable.

<sup>17</sup> There are, in fact, several versions of this hypothesis. See Lieber and Scalise (2007).

<sup>18</sup> Inclusivity in catena-based dependency grammar: a unit of form U is more inclusive than another unit V, if more node combinations in a given expression qualify as U-type units, than as V-type units. There are three (four, if we include the suffix) constituents in each (9a) and (9b). The text lists six catenae obtained on a catena-based analysis. Since there are three (or four) constituents in (9a) and (9b), but six catenae in (9c), the catena is more inclusive than the constituent. Conversely, constituents are more exclusive than catenae.

<sup>19</sup> This would also be an important argument against the structuralist piece-based approach (Bloomfield 1933; Harris 1942; Hockett 1947, 1954; Nida 1948) because this approach is constituent-based.

One should, however, not overestimate the reach of word/lexeme/paradigm-based morphology. The following authors also subscribe to internal word structure: Sadock (1991), Di Sciullo (2005), Williams (2011), and Distributed Morphology (Halle & Marantz 1993; Harley & Noyer 2003; Embick & Noyer 2001, 2007; Embick 2003; and others).<sup>20</sup>

A final comment addresses the cognitive and psycholinguistic implications of the catena. There is evidence from syntax, in particular from phenomena such as displacement, ellipsis, idioms, and constructions, that catenae are the primary unit of syntactic structure (Groß & Osborne 2009; Osborne & Groß 2012a; Osborne, Putnam & Groß 2012). Catena-based dependency grammar (Osborne & Groß 2012a; Groß & Osborne 2013) has in common with Cognitive Grammar and Construction Grammar the assumption of continua (syntax-morphology, grammar-lexicon, free-bound, etc.). A further commonality lies in rejecting specialized cognitive modules, instead preferring to assume a general-purpose module. The lexicon contains catenae (or, as Nanosyntax suggests, “subtrees”, see Starke 2009: 2), and these are acquired by exposure, and fortified by repetition. As such catena-based dependency grammar concurs also with Usage- and Frequency-based accounts (Bybee 2003, 2010). However, the assumption of a general-purpose module, and the adoption of usage- and frequency-based principles do not logically imply that the units involved in the storing and processing of language must be words.

### 3. The parts of the desiderative construction

This section intends to clear the way for the analysis of the Hebrew desiderative construction. The first subsection gives a brief introduction into the root-and-pattern morphology of Hebrew, thereby pointing out the challenges for piece-based accounts of morphology, and morphosyntax. The second subsection introduces the future tense formation in Hebrew, which appears in the desiderative construction. The final subsection briefly addresses cliticization because one unit in the desiderative construction is a clitic.

---

<sup>20</sup> The reservation against Distributed Morphology is that it utilizes movement. The motivation for movement stems from overcoming bracketing paradoxes that occur whenever displacement is analyzed with constituent structure.

### 3.1 Transfixes and radicals

Hebrew is a root-and-pattern language. In such languages, non-concatenative phenomena play a more prominent role than in inflecting, or agglutinating languages. This account follows McCarthy (1981) in viewing radicals (used instead of “root”), and patterns as distinct meaning-bearing units. This decision stems from the possibility that a piece-based representation of morphological phenomena is in principle possible because catena-based dependency morphology can represent structure in the vertical, i.e. dominance, dimension (cf. section 2.3). Assuming radicals is, however, not universally the case in Hebrew linguistics. Bat-El (2001: 13) gives a brief overview over proponents, opponents, and linguists who ignore the issue.

The challenge to piece-based morphological theories that languages such as Hebrew posit, is demonstrated now with several possibility expressions from Modern Hebrew. Consider the next examples:

- (11) a. ʔaxil                      b. savir                      c. naʔil                      d. patir  
       ‘edible’                      ‘reasonable’                      ‘can be locked’                      ‘solvable’

In examples (11a–d), the underlined letters designate a root phoneme, and the remaining letters a pattern phoneme. In (11a), for instance, the root phonemes are /ʔ/, /x/, and /l/, which together form the radical *ʔxl*, meaning ‘eat’. In (11b), the radical is *svr*, meaning ‘reason’. In (11c), the radical is *nʔl*, which means ‘lock in’. Finally, the radical *ptr* in (8d) means ‘solve’.<sup>21</sup>

The challenge these examples posit is evident. If the phonemes /ʔ/, /x/, and /l/ in (11a) are an expression of the meaning “eat”, then this expression must include the possibility expression because the vowels of the transfix (= pattern) appear between the root consonants. Since all examples in (11) express possibility, and since all examples in (11) contain the vowels /a/ and /i/, one may well argue that the vocalic transfix  $\bar{a}\bar{i}$  is the evident candidate to which to assign the possibility meaning.<sup>22,23</sup> However, this transfix (and others like it) cannot constitute a unit of form in any theory of

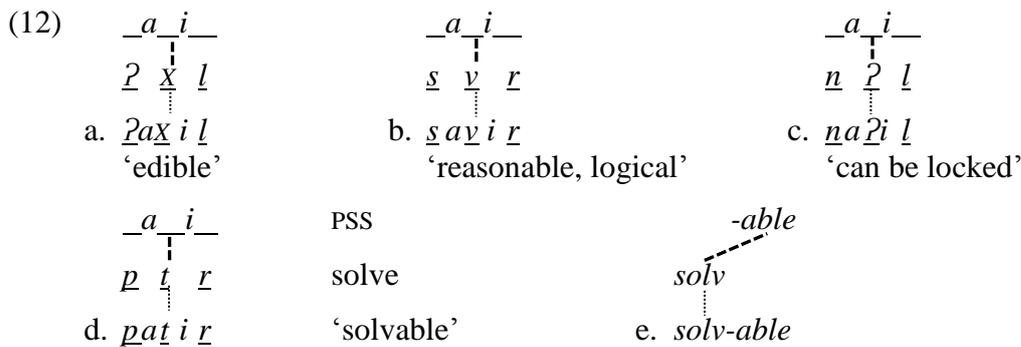
<sup>21</sup> One reviewer criticizes the usage of verbal meanings. I agree that root meanings are probably more abstract. The choice here is merely a matter of convenience.

<sup>22</sup> In the text, a transfix is shown as  $\bar{transfix}$ .

<sup>23</sup> One reviewer points to examples that contain the /a-i/ pattern, but fail to express possibility. But that does not impact the fact that this pattern is highly productive in the verbal *PA'AL* and *PI'EL* classes as the expression of possibility.

morphology that cannot isolate units in the vertical dimension. The radicals, too, may be seen as expressions of their respective meanings. The same problem applies here, namely that these phonemes fail to form a string, i.e. a unit of form that is continuous with respect to the horizontal dimension.

If a theory of morphology can isolate the vertical from the horizontal dimension of representation, as the catena allows one to do, then it becomes possible to separate the transfixes from the radicals in a meaningful way. In the examples (11a–d), the transfixes dominate their radicals, because the entire expressions distribute like adjectives marked for possibility expressions, rather than like verbs. In the current dependency-morphological account the tree structures of the examples (11a–d) are represented in the following manner:



The structures in (12a–e) show representations of the meaningful units making up the individual words. One example, (12d), is explained in detail, the examples (12a–c) are constructed in the same fashion. In (12d), the tree structure distinguishes two units: the transfix on top, and the radical immediately below it. The transfix dominates the radical, and this fact is represented by the vertical dotted edge linking the two units. The underbars help to identify the positions into which the root consonants are inserted; the appearance of the consonants depends on the transfix, hence the underbars are part of the transfix.<sup>24</sup>

The transfix dominates the radical because the word in its entirety behaves like an expression marked for possibility. The possibility transfix is derivational; it produces an adjective. All the examples in (12) behave like adjectives. This is the kind of argument that also applies to English

<sup>24</sup> This approach is reminiscent of Autosegmental Phonology (Goldsmith 1976, 1990; McCarthy 1981; Lieber 1987), which can be viewed as an attempt to introduce a vertical dimension into the analysis.

*solvable*, example (12e). Every occurrence of “X-able” is an adjective in English, hence the suffix should dominate the verb there, too.<sup>25</sup> Since the English possibility expression is a suffix, rather than a transfix, it can be separated from the lexical unit not only in the vertical, but also in the horizontal dimension. For this reason the dotted edge in (12e) is slanted.

The novelty in (12a–d) is the separation of meaningful units in the vertical, i.e. dominance, dimension, rather than in the horizontal, i.e. precedence, dimension. Looking exclusively at the vertical structure of (12a–d), the assumption that the radical, and the transfix, form units of form is justified because nothing intervenes between the two units in the vertical dimension. Abstracting from the horizontal dimension makes it possible to represent the gloss in a vertical fashion. The vertical gloss in (12d) is thus more informative than a horizontal gloss (such as e.g. *solve.PSS*). The comparison of (12d) with its English equivalent (12e) shows that a representation that identifies units of meaning and units of form across languages with fundamentally different word structure is achievable.

### 3.2 Future tense

The introduction into root-and-pattern morphology above has been brief, but was necessary because the desiderative construction requires that the verb in the subordinate clause be in the future tense.<sup>26</sup> The MH future tense serves two purposes: it is used to express events that have not happened yet, i.e. irrealis and imperfective cases. But it is also used as the imperative. A number of verbs with irregular features continue to form their Standard Hebrew imperative forms.

According to Bat-El (1994: 582), MH verbs have five different conjugations, so-called *binyan*. However, the introduction below will be concerned only with one conjugation, the so-called *PA'AL* conjugation. The example radical will be *ft̥r*, the radical of example (12d).

---

<sup>25</sup> That includes cases like *convertible* which have undergone conversion after ellipsis, as they derive from *convertible car/automobile*.

<sup>26</sup> The term “future tense” is used here only as a label, and should not be taken as representing the grammatical meaning associated with this tense. Gesenius (1909: 117), for instance, named this tense as the imperfective, rather than the future. I refrain from a discussion of the appropriateness of this or other distinctions, and continue to use the term “future tense” as a label of convenience. But see also Coffin and Bolozky (2005: 38–40).

Future tense forms and imperatives share the property that they both appear with vowel transfixes ( $\bar{o}$  or  $\bar{a}$ ), or the epenthetic vowel /e/. Future tense and imperative expressions are thus marked as irrealis or imperfectives. Bat-El (2002) argues that MH imperatives are formed by true truncation from future tense forms.

The future tense is expressed by prefixes. The first person is expressed by the prefix *ʔe-* in the singular, and *ni-* in the plural. The second person prefix, and the third person singular feminine prefix is *ti-*. The third person is expressed with the prefix *ji-*.

In addition, suffixes appear. In the second person singular feminine the suffix *-i* appears, and in the second and third person plural the suffix *-u* is used. The next examples with the radical *ʔtr* illustrate the future tense forms:

- |   |  |   |
|---|--|---|
| (13) a. <i>ʔe-ʔtr</i><br>'I will solve'         |  | b. <i>ni-ʔtr</i><br>'we will solve'         |
| (14) a. <i>ti-ʔtr</i><br>'you[M.SG] will solve' | a'. <i>ti-ʔter-i</i><br>'you[F.SG] will solve' | b. <i>ti-ʔter-u</i><br>'you[PL] will solve' |
| (15) a. <i>ji-ʔtr</i><br>'he will solve'        | a'. <i>ti-ʔtr</i><br>'she will solve'          | b. <i>ji-ʔter-u</i><br>'they will solve'    |

The examples in (13) show the first person, (14) shows the second person, and (15) the third person. (a)-examples show the singular form (13a), or the masculine singular forms (14a, 15a). The barred examples show feminine singular forms. The (b)-examples show plural forms.

The vowel /o/ in (13a and 13b), (14a), and (15a, a') is the transfix  $\bar{o}$  marking irrealis or imperfective. The vowel /e/ in (14a' and 14b) and (15) is viewed as epenthetic. I will part from this conventional assumption, and assume that this vowel is a suppletive transfix  $\bar{e}$  that appears whenever a suffix appears.

The forms (14a, 14a' and 14b) can also be used as imperatives in MH, but truncated forms of (14a, 14a' and 14b) are also possible:<sup>27</sup>

- |  |                                      |                                    |
|--|--------------------------------------|------------------------------------|
| (16) a. <i>ʔtr!</i><br>'solve!' [SG.M] | b. <i>ʔter-i!</i><br>'solve!' [SG.F] | c. <i>ʔter-u!</i><br>'solve!' [PL] |
|--|--------------------------------------|------------------------------------|

<sup>27</sup> I am grateful to an anonymous reviewer for pointing me toward Bat-El's (2002) truncation analysis.

The forms in (16) will be analyzed first, and then the dependency morphological representation of future tense forms will follow. The suffixes are expressions of number: *-i* expresses the singular, and *-u* the plural. They are used whenever the prefix does not conclusively express number.

(17)	$\begin{array}{c} \text{---}o\text{---} \\   \\ \underline{ft} \quad \underline{r} \\ \vdots \\ \underline{ft} \quad \underline{or} \\ \text{'solve!' [M.SG]} \end{array}$	IRR solve	$\begin{array}{c} \text{---}e\text{---} \\ \diagup \text{---}i \\   \\ \underline{ft} \quad \underline{r} \\ \vdots \\ \underline{ft} \quad e \quad \underline{r-i} \\ \text{'solve!' [F.SG]} \end{array}$	F.SG IRR <sup>s</sup> solve	$\begin{array}{c} \text{---}e\text{---} \\ \diagup \text{---}u \\   \\ \underline{ft} \quad \underline{r} \\ \vdots \\ \underline{ft} \quad e \quad \underline{r-u} \\ \text{'solve!' [PL]} \end{array}$	PL IRR <sup>s</sup> solve
a.			b.		c.	

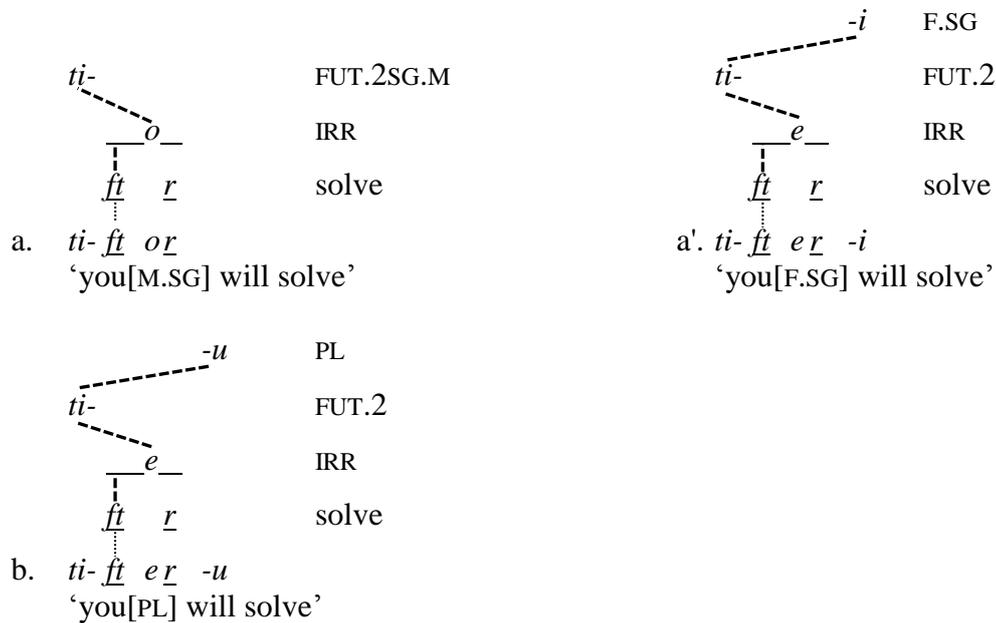
In (17), IRR is used as a label to reference irrealis **or** imperfective grammatical meaning. The superscript in (17b and 17c) stands for SUPPLETIVE. Expressions of number appear farther from the stem than expressions of mood. Hence, one should expect the number suffixes to dominate the radical (here indirectly), and the transfixes. Mood should dominate lexical material, and this is how all the examples in (17) are represented.

Having represented the imperative forms, the analysis now proceeds to the future tense forms.<sup>28</sup> The tree representations below follow the format used in the examples (13–15), repeated below as (18–20):

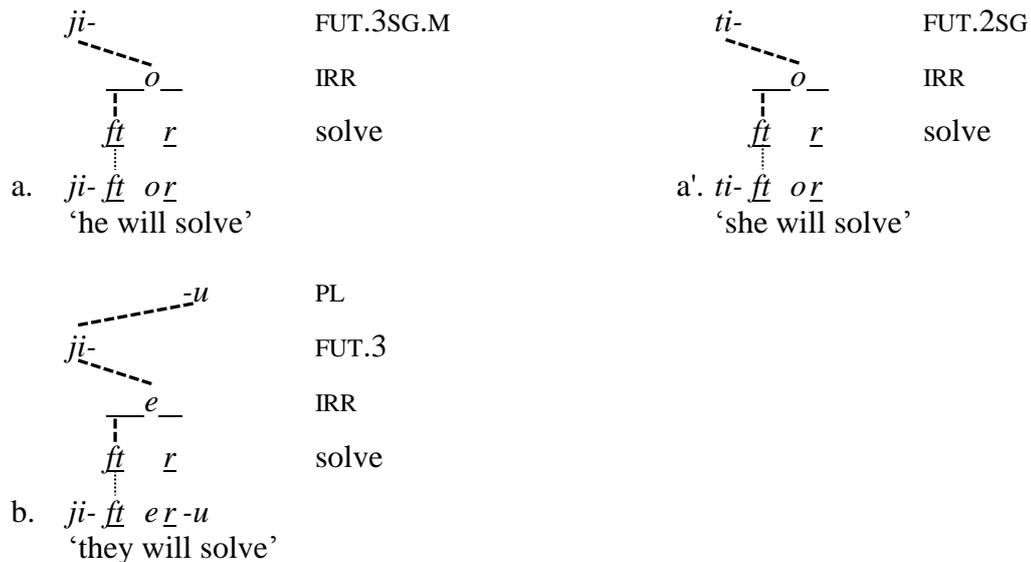
(18)	$\begin{array}{c} \text{?}e\text{---} \\ \diagdown \text{---}o\text{---} \\   \\ \underline{ft} \quad \underline{r} \\ \vdots \\ \text{?}e\text{---} \underline{ft} \quad \underline{or} \\ \text{'I will solve'} $	FUT.1SG IRR solve	$\begin{array}{c} ni\text{---} \\ \diagdown \text{---}o\text{---} \\   \\ \underline{ft} \quad \underline{r} \\ \vdots \\ ni\text{---} \underline{ft} \quad \underline{or} \\ \text{'we will solve'} $	FUT.1PL IRR solve
a.			b.	

<sup>28</sup> The presentation here is limited to the default of the PA'AL conjugation. The vowels of the prefixes are seen here, for the purpose of simplification, as parts of the prefixes. Yet, these vowels can change: the verb may be irregular; the root may have a guttural or laryngeal as the initial consonant, etc. These issues are neglected here.

(19)



(20)



All structures in (17–20) follow the assumption by Bybee (1985) that there is an order in which affixes appear with respect to the lexical core. Her basic assumption is that the category valency is realized closer to the verb, than the category mood, for instance. The categories farthest from the lexical core are person and number, whereby number follows person.<sup>29</sup>

<sup>29</sup> The current account subscribes to almost all of Bybee's (1985) claims about affix order. It differs in one very important respect, though. Affix order is not seen as operating in the horizontal dimension, but in the vertical dimension. Bybee is concerned

(19a' and 19b) and (20b) reflect that assumption by viewing the number suffixes dominating the person/tense prefixes. The assumption here is that number dominates tense, which dominates mood. Future tense in Modern Hebrew is an instance of “cumulative exponence” (Matthews 1991: 179) because the respective prefixes usually serve to express both tense, and at least person and number, sometimes also genus. When the gloss in (18–20) only references person, but not genus or number, it means that the default is set, which is interpreted as the masculine singular, in the absence of further grammatical markers.

The structures above are now discussed in detail. Example (18a) shows the first person singular of the future tense. In Modern Hebrew, the first person is not sensitive to genus. The tense/person prefix dominates the unmarked imperative transfix, which dominates the radical. (18b) shows the first person plural.

Example (19a) is not unambiguous. Without a subject, it is addressed to a male person. Together with a subject with feminine genus, it expresses the third person singular feminine (20a'). In (19a'), the suffix *-i* marks the expression as second person feminine singular. The presence of the suffix blocks the third person reading of the prefix *ti-*. The catena *ti-...-i* expresses the future tense for the second person singular feminine because, in combination, all relevant properties are present.<sup>30</sup> Example (20a) shows the third person default form, which expresses the singular masculine.

### 3.3 Cliticization

A clitic is a prosodically dependent, but syntactically free unit. “Prosodically dependent” means that a unit cannot project its own prosodic word structure, and is therefore integrated into the prosodic word structure of another unit. Free morphs are, generally, syntactically free, and project their prosodic word structure. In compounds, compound parts relinquish these features. Affixes are never prosodically or syntactically independent. Clitics may be part of a prosodic word, and yet not maintain any kind of dependency relationship with any of the other units within this prosodic word. The introduction is necessary because the Hebrew complementizer

---

with serial order, rather than with dominance order. It turns out, though, that dominance, rather than precedence, is the critical dimension.

<sup>30</sup> The notation *ti-...-i* is the convention of referencing catenae within the text. This particular example means that the prefix *ti-* and the suffix *-i* form a catena that is interrupted by other material in the horizontal dimension.

*še=* is a clitic. It will be kept brief, though, since the issue of clitics has already been addressed within the current framework (Groß 2011b, 2014).<sup>31</sup> An English example is used here to first show that clitics posit problems for constituent-based grammars:

- (21) a. *the Queen of England's hat*  
 b. [*the*] [*Queen*] [*of*] [*England's*] [*hat*]      prosodic word structure  
 c. [[[[*the*] *Queen* [*of*] [*England*]]]'s] *hat*]      syntactic constituent structure

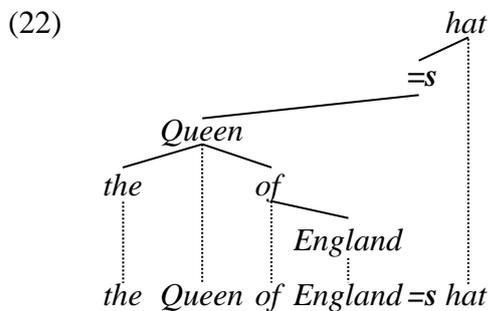
Example (21a) contains the English possessive, which is a clitic (in bold script). (21b) shows the prosodic word structure of (21a). The possessive is clearly part of the word structure projected by *England*. (21c) shows the constituent structure of (21a). Here, the clitic is not part of *England*, but rather of *the Queen of England*. The reason for this is that even though as (21b) shows, it is part of the prosodic word structure of *England*, the clitic does not entertain an immediate syntactic relationship with its host. In other words, the clitic does not establish a relationship between *England* and *hat*, but rather between *Queen* and *hat*. Since *Queen* is the projecting node of the NP *the Queen of England*, this entire NP must be the constituent with which the clitic has a structural relationship.<sup>32</sup>

The structures (21b) and (21c) are at odds because they cannot be both true at the same time. The underlying cause for this paradox is the notion of the constituent because this notion references the horizontal dimension and the vertical dimension simultaneously. This means that constituent-based grammars are forced to make a stronger claim than necessary: the possessive clitic must dominate the NP *the Queen of England*, but also be part of the most subordinated unit in this NP, namely *England*. The attempt to view the phenomenon through the lens of constituent structure makes cliticization look more difficult than it really is.

<sup>31</sup> For reasons of brevity, the examples in this section are kept simple. There is, however, more to cliticization. See Groß (2011b, 2014) for a detailed discussion of various languages, including K<sup>w</sup>ak<sup>w</sup>'ala (Anderson 2005: 16), Wackernagel clitics, clitic climbing, clitic doubling, etc.

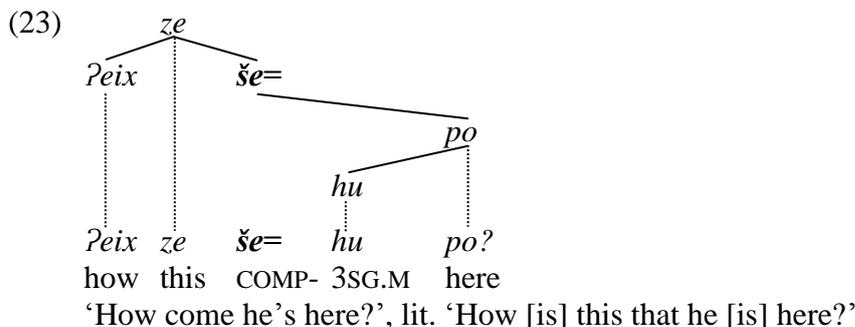
<sup>32</sup> Word-grammatical dependency grammars are not immune to the challenges posited by clitics. Cliticization challenges, in particular, the widely acknowledged assumption that a node may not be immediately dominated by two or more nodes (Mel'čuk 1988: 23). For (21a), this means that the word node *England's* would have to be a dependent of the nodes *of* and *hat* simultaneously.

In a catena-based approach, the necessity to conflate the horizontal (= prosodic) dimension and the vertical (= dominance) dimension is absent, and hence one can avoid overly strong claims. The claim about prosodic structure (horizontal dimension) is identical to (21b): the clitic is integrated into the prosodic word structure of *England*. This is shown below as a hyphen on the clitic directed toward the prosodic host, and through the absence of a projection edge on the clitic. The claim about dominance only identifies the two nodes to which the clitic connects in the vertical dimension. The next representation shows the example (21) as a dependency structure:



Example (22) shows the dependency structure of (21a). In the horizontal dimension, the clitic is integrated into the prosodic word structure of the word immediately to its left, i.e. *England*. It need not also entertain an immediate dominance relationship (vertical dimension) with this word. Rather, in the vertical dimension, the clitic is immediately dominated by *hat*, and it immediately dominates *Queen*. This makes sense because the clitic establishes a possessive relationship between these two nodes. Cliticization is, thus, a phenomenon in which the two dimensions are split, and fulfill two entirely different roles.

Cliticization occurs with the Hebrew complementizer *še=*, appearing in the desiderative construction. Consider the next Hebrew example:



Since Hebrew does not know copulae (in the present tense), *ze* ‘this’ is seen as the main clause root, and *po* ‘here’ as the predicate of the subordinated clause. The complementizer clitic *še=* (bold) is part of the prosodic word structure of the subject *hu* ‘he’ of the subordinated clause. However the clitic does not entertain an immediate dependency relationship with its host, rather it immediately dominates *po*, the predicate of the subordinated clause. Using the notions established in section 2.1, *še=hu* is a string, but not a catena. On the other hand, *še=...po* is a catena.

#### 4. The Hebrew desiderative construction

This section is, finally, concerned with the Hebrew desiderative. In the previous section, it has been shown how transfixes are represented in this account. A sketch of the Hebrew future tense has also been provided. Finally, cliticization was briefly addressed. All these issues are important in the description of the Hebrew desiderative.

First, the construction to be discussed below is confined to a specific desiderative meaning, namely that of the main clause subject wishing others to do something.<sup>33</sup> This kind of construction is transclausal in the sense that a verb of volition dominates an obligatory complementizer. German, too, expresses desideratives using a complementizer (in boldface):

- (24) *Ich will, dass er kommt.*  
 1SG VOL COMP 3SG.M come.3SG.NPST  
 ‘I want him to come.’

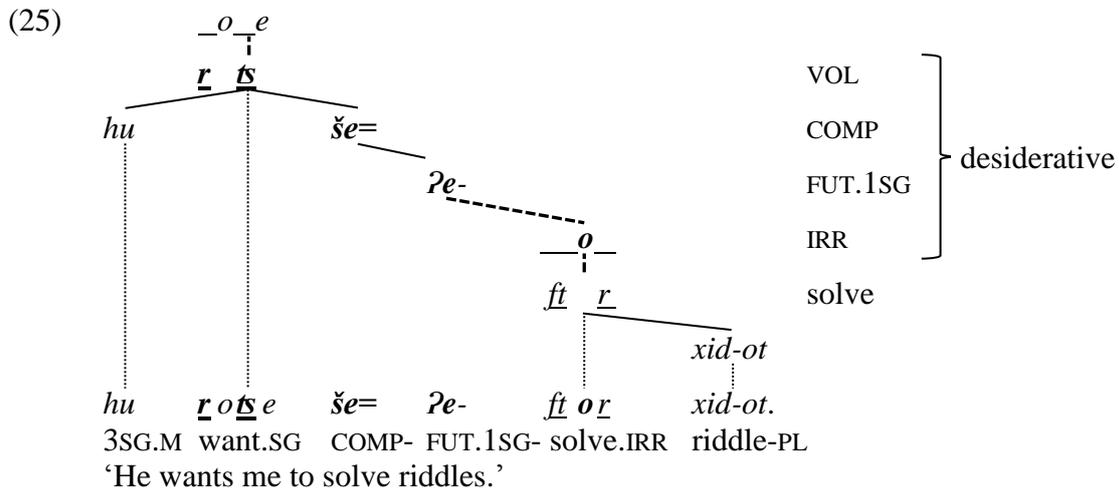
Within the subordinated clause, certain grammatical meanings, namely future tense, and person, number, and possibly genus must be expressed. The desiderative meaning in the strict sense is restricted to the following verbs: *ratsa* ‘want’, *bikef* ‘wish’, *ixel* ‘wish well’, *kiva* ‘hope’, as well as several others.<sup>34</sup> In their desiderative meaning, these verbs require that four conditions be met: 1. a subordinate clause headed by the clitic *še=* ‘that’ must be present. 2. The future tense must be present in the subordinate clause. 3. Future tense prefixes have cumulative exponence. The one exception is the prefix *ti-*, which can, in the absence of a genus or number

<sup>33</sup> In Palmer (2001: 131–135), the notion of desiderative includes all expressions of desiring.

<sup>34</sup> The verbs are given here in the past tense 3SG.M form, because that is used as the lemma form in dictionaries.

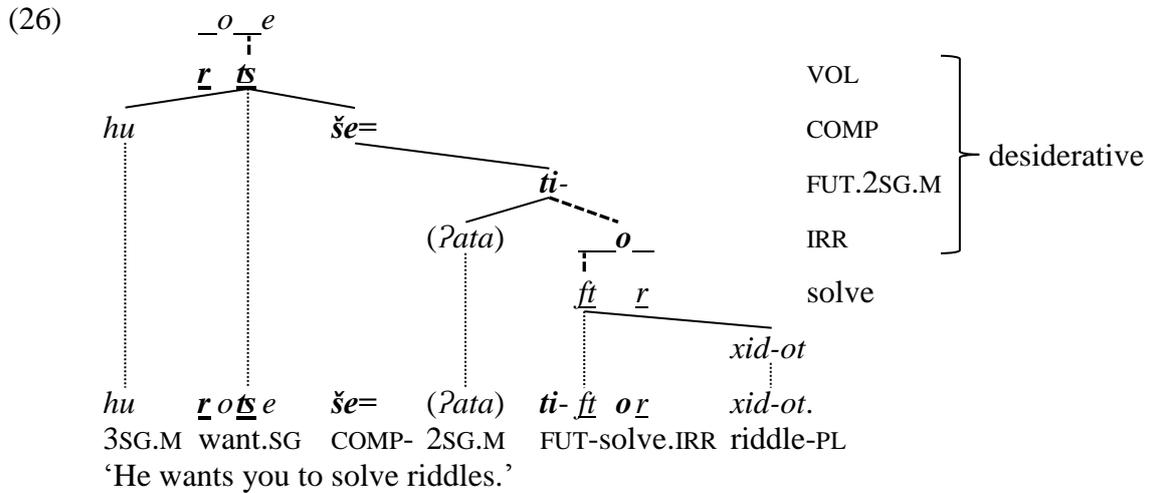
suffix, express the second person singular masculine, or the third person singular feminine. This prefix is underspecified with respect to person, genus, and number. In the second person, the optionality of the subject is licensed by the context, but in the third person, a subject (pronoun) is usually not omitted. If the subject is a pronoun, then this pronoun is deemed part of the desiderative construction because it serves to fully specify the grammatical information in the subordinated clause. 4. The units of form that contribute to the constructional meaning of the desiderative must qualify as a catena.

In the examples below, the verb *ratsa* ‘want’ will be used as the verb of volition. The verb in the subordinated clause remains *ft* ‘solve’. Consider the next sentence:



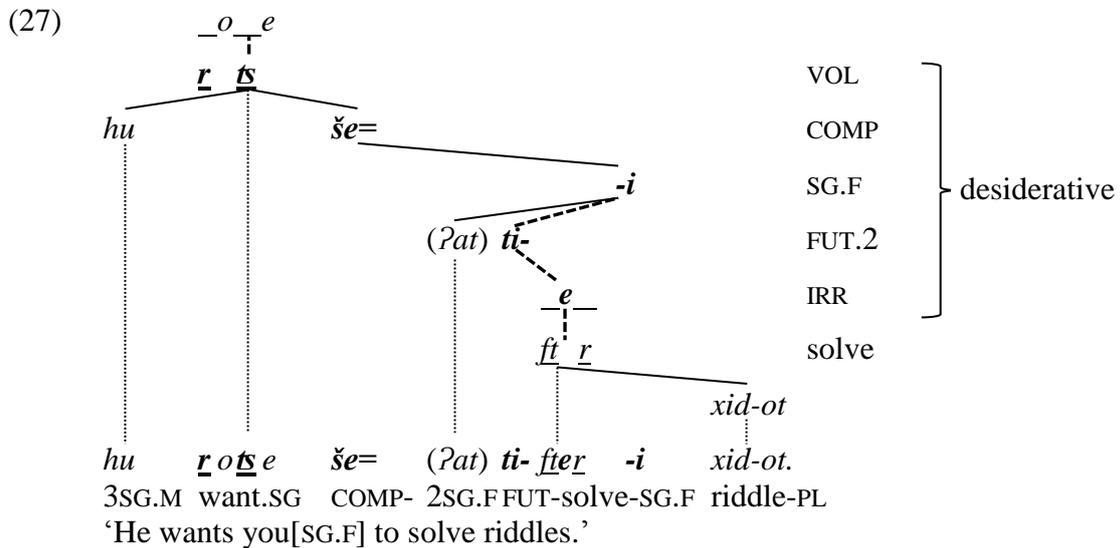
The units in boldface are those involved in the desiderative construction. As before, root consonants are underlined. All four conditions are satisfied: 1. A subclause headed by *še=* is present. 2. Future tense is present (*ʔe-*). 3. Associated person and number properties are specified (*ʔe-* = 1SG); genus properties are not specified in the first person. 4. The units involved qualify as a catena.

With the verb in the subordinate clause in the second person, a pronoun is optional, if the subject is masculine. Optionality is expressed by round brackets:



The bracketed expression *?ata* ‘2SG.M’ is facultative.<sup>35</sup> If the pronoun is viewed as part of the desiderative construction, even though it is facultative, the fourth condition is still satisfied: the units involved qualify as a catena. If the pronoun is present, it forms the host for the clitic *še=*. If it is absent, the verb *ti-ftor* becomes the host.

When the prefix *ti-* appears together with the suffix *-i* then the third condition is still met:

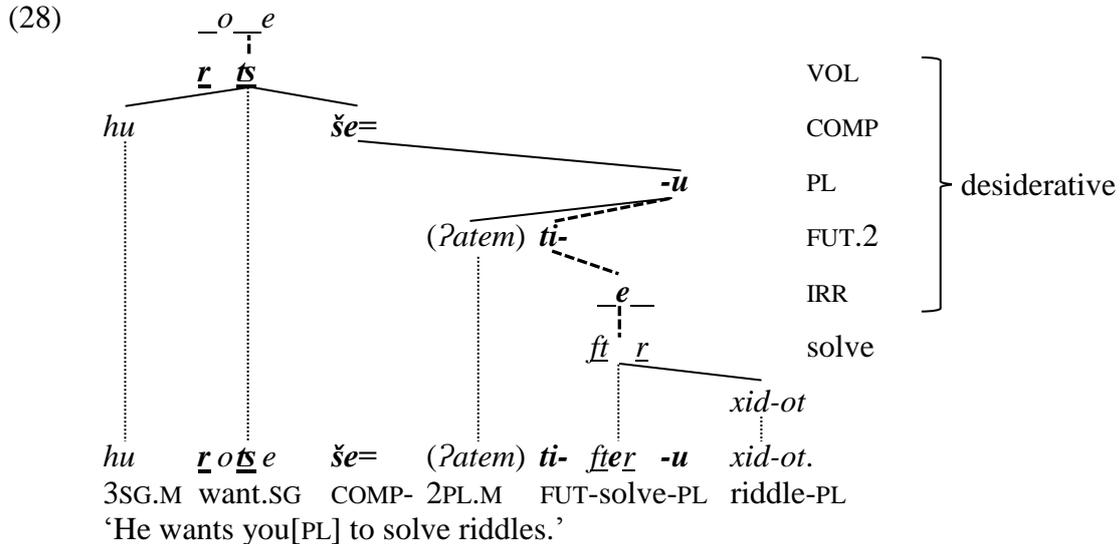


In the second person singular feminine, the suffix *-i* must be present. Hence the appearance of a subject pronoun is optional. Note that this analysis

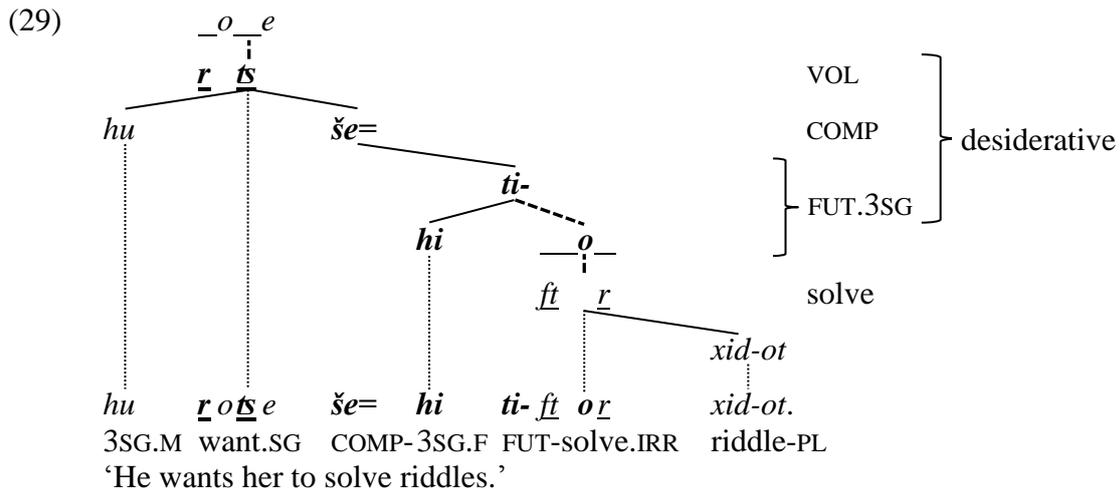
<sup>35</sup> The subject pronoun is seen here as a dependent of the tense prefix. Doing so corresponds to the assumption of a TP in, e.g. minimalist grammars.

continues to capture the constructional units as a catena, i.e. the fourth condition is met.

A similar constellation appears in the plural:



Here, the pronoun *patem* ‘you.PL’ is facultative. The plural suffix is sufficient in specifying number.<sup>36</sup> If the subject in the subordinated clause is in the third person, a pronoun tends to appear. In case of the feminine third person singular, the presence of a pronoun results in full specification:



<sup>36</sup> Note, though, that *patem* can express the unmarked second person plural, or the second person plural masculine. The feminine version is *paten*.

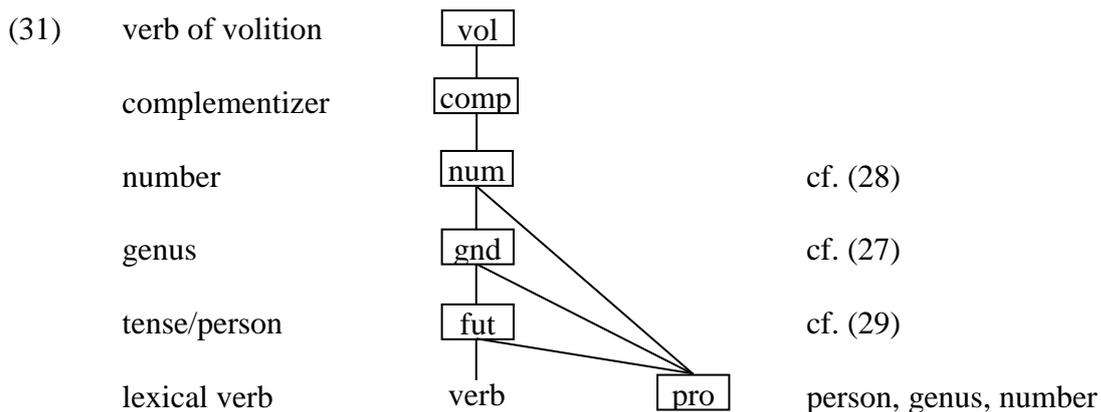
The feminine pronoun in (29) is in boldface (as part of the construction). Note that all four conditions are satisfied. A third person singular masculine pronoun tends to appear as the subject in the subordinate clause, even though full specification is accomplished by the prefix.

The third person plural forms of verbs function as impersonal verb forms, but do then appear without the third person plural pronoun *hem* ‘they’. In this case the absence of this specific pronoun is the indication that the verb is impersonal:

- (30) a. *ʔani rotse še= **hem** ye-dʔ-*u*.*  
 1SG want.NPST COMP- 3PL FUT.3-know-PL  
 ‘I want them to know.’
- b. *ʔani rotse še= ye-dʔ-*u*.*  
 1SG want.NPST COMP- FUT.3-know-PL  
 ‘I want it to be known.’

In (30a), the boldface pronoun *hem* appears in the subordinate clause. Hence the sentence is understood as referencing a personal subject of the verb *ye-dʔ-*u**. On the other hand, (30b) lacks a pronoun, and since the verb in the subordinate clause is in the third person plural, the interpretation of an impersonal verb is warranted.<sup>37</sup>

In summary, the desiderative construction can be captured in a purely vertical and abstract fashion:



The verb of volition, here *ratsa* ‘want’, is the top node of the construction. Clearly, this verb may be marked by any grammatical form that does not overrule its valency potential, and still be able to participate in the

<sup>37</sup> I am grateful to an anonymous reviewer for bringing this issue to my attention.

construction. It immediately dominates the complementizer. The complementizer dominates material that appears at the top of clauses, namely units that express tense, mood, number, genus, etc. The highest units below the complementizer are the pure number suffixes (cf. ex 28). If such a suffix is present, a co-occurring pronoun must depend on it. The number/genus suffix is also possible (cf. ex 27). If it is present, the pronoun must depend on it. Next are tense/number/person prefixes (cf. ex 25 and 26). The person features are usually expressed together with tense, i.e. as a prefix. When a pronoun appears, it must depend on the prefix. Whenever third person subject pronouns appear, they are considered as part of the construction.

## 5. Summary

This paper has attempted to capture the Hebrew desiderative, a transclausal periphrastic construction within catena-based dependency morphology. Section 2 briefly informed the reader on dependency grammatical notions and terms. Section 2.1 introduced the central term of this account, namely the catena. Section 2.2 included a brief history of dependency morphology. Section 2.3 introduced central notions of catena-based morphology, and it showed how the catena is used in morphology and morphosyntax. In section 2.4 principal reservations against other approaches to morphology were voiced.

The individual units involved in the desiderative construction posit certain challenges to theories of morphology. In order to develop an understanding of catena-based dependency morphology, section 3 touched on three issues: section 3.1 addressed transfixes and radicals, section 3.2 introduced the future tense paradigm, and section 3.3 briefly addressed cliticization. This was necessary because certain transfixes must be viewed as residing outside of the construction, for instance tense or other markers on the construction root, the verb of volition. On the other hand, certain other trans-, pre-, and suffixes must be viewed as constructional units. Further, the verb within the subordinated clause must be marked with future tense prefixes. Finally, the complementizer is a clitic, which necessitated a brief layout how cliticization works in dependency morphology.

Section 4 illustrated with several examples how the construction can be represented within the proposed framework, and that the units of the desiderative construction always qualify as a catena. Further, it showed that

the desiderative construction is neither word-based, nor constituent-based. Rather the construction cuts into words, namely it reaches into the verb of volition, of which only the radical participates in the construction. At the lower end of the construction, the morphs that together express future tense, i.e. the prefix, and the transfix must be part of the construction, while the lexical verb which they dominate is excluded. Thus, the current account makes the case for an entirely surface-based analysis of periphrastic constructions, taking the Hebrew desiderative as its concrete example.

## References

- Ágel, Vilmos; Eichinger, Ludwig; Eroms, Hans-Werner; Hellwig, Peter; Heringer, Hans Jürgen & Lobin, Henning (2003) *Handbücher zur Sprach- und Kommunikationswissenschaft: Dependenz und Valenz*, Vol.1. Berlin/New York: Walter de Gruyter.
- (2006) *Handbücher zur Sprach- und Kommunikationswissenschaft: Dependenz und Valenz*, Vol.2. Berlin/New York: Walter de Gruyter.
- Ágel, Vilmos & Fischer, Klaus (2010) 50 Jahre Valenztheorie und Dependenzgrammatik. *Zeitschrift für germanistische Linguistik* 16: 249–290.
- Anderson, John (1980) Towards dependency morphology: The structure of the Basque verb. In John Anderson & Colin J. Ewen (eds.), *Studies in Dependency Phonology*, pp. 221–271. Ludwigsbuurg: R.O.U. Strauch.
- Anderson, Stephen R. (1992) *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- (2005) *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.
- Aronoff, Mark (1976) *Word Formation in Generative Grammar*. Cambridge, Massachusetts: MIT Press.
- (1992) Stems in Latin verbal morphology. In Mark Aronoff (ed.), *Morphology Now*, pp. 5–32. Albany: State University of New York Press.
- Bat-El, Outi (1994) Stem classification and cluster transfer in Modern Hebrew. *Natural Language and Linguistic Theory* 12: 571–596.
- (2001) In search for the roots of the C-root: The essence of Semitic morphology. Workshop on Root and Template Morphology, USC: Los Angeles. May 6, 2001. <<http://www.outibatel.com/wp-content/uploads/2010/12/Bat-El-2001-C-root.pdf>> (25 July 2014)
- (2002) True truncation in colloquial Hebrew imperatives. *Language* 78 (4): 651–683.
- Bloomfield, Leonard (1933) *Language*. New York: Holt.
- Booij, Geert (2010) *Construction Morphology*. Oxford/New York: Oxford University Press.
- Bybee, Joan L. (1985) *Morphology: A Study of the Relation Between Meaning and Form*. Amsterdam: John Benjamins Publishing Company.

- (2003) *Phonology and Language Use*. Revised edition. Cambridge: Cambridge University Press.
- (2010) *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Coffin, Edna Amir & Bolozky, Shmuel (2005) *A Reference Grammar of Modern Hebrew*. Cambridge: Cambridge University Press.
- Creider, Chet & Hudson, Richard (1999) Inflectional morphology in Word Grammar. *Lingua* 107: 163–187.
- Di Sciullo, Anna Maria (2005) *Asymmetry in Morphology*. Cambridge: MIT Press.
- Embick, David (2003) Linearization and local dislocation: Derivational mechanics and interactions. *Linguistic Analysis* 33 (3–4): 303–336.
- Embick, David & Noyer, Rolf (2001) Movement operations after syntax. *Linguistic Inquiry* 32 (4): 555–595.
- (2007) Distributed Morphology and the syntax/morphology interface. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces*, pp. 289–324. Oxford: Oxford University Press.
- Engel, Ulrich (1994) *Syntax der deutschen Gegenwartssprache*. 3rd edition. Berlin: Erich Schmidt.
- Eroms, Hans-Werner (1985) Eine reine Dependenzgrammatik für das Deutsche. *Deutsche Sprache* 13: 306–326.
- (2000) *Syntax der deutschen Sprache*. Berlin: Walter de Gruyter.
- Gesenius, Friedrich Heinrich Wilhelm (1909) *Gesenius' Hebrew Grammar, as Edited and Enlarged by the Late E. Kautzsch. Second English Edition, Revised in Accordance with the Twenty-Eighth German Edition*. Translated by A.E. Cowley. Unabridged Republication. 2006. Mineola, NY: Dover Publications Inc.
- Goldsmith, John (1976) *Autosegmental Phonology*. Doctoral dissertation. MIT.
- (1990) *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell.
- Groß, Thomas (1999) *Theoretical Foundations of Dependency Syntax*. München: iudicium.
- (2010) Chains in syntax and morphology. In Otaguro Ryo, Kiyoshi Ishikawa, Hiroshi Uemoto, Kei Yoshimoto & Yasunari Harada (eds.), *Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation at Tohoku University*, pp. 143–152. Tokyo: Waseda University.
- (2011a) Catenae in morphology. In Kim Gerdes, Eva Hajičová & Leo Wanner (eds.), *Depling 2011*, pp. 47–57, Barcelona: Pompeu Fabra University. <<http://depling.org/proceedingsDepling2011/papers/gross1.pdf>> (24 August 2014).
- (2011b) Clitics in dependency morphology. In Kim Gerdes, Eva Hajičová & Leo Wanner (eds.), *Depling 2011*, pp. 58–68. Barcelona: Pompeu Fabra University. <<http://depling.org/proceedingsDepling2011/papers/gross2.pdf>> (24 August 2014).
- (2011c) Transformational grammarians and other paradoxes. In Igor Boguslavsky & Leo Wanner (eds.), *MTT 2011*: 88–97. Barcelona: Pompeu Fabra University. <<http://meaningtext.net/mtt2011/proceedings/papers/Gross.pdf>> (24 August 2014).
- (2014) Clitics in dependency morphology. In: Kim Gerdes, Eva Hajičová & Leo Wanner (eds.), *Dependency Linguistics: Recent Advances in Linguistic Theory*

- Using Dependency Structures*, pp. 229–252, Amsterdam/New York: John Benjamins.
- Groß, Thomas & Osborne, Timothy (2009) Toward a practical dependency grammar theory of discontinuities. *SKY Journal of Linguistics* 22: 43–90.
- (2013) Katena und Konstruktion: Ein Vorschlag zu einer dependenziellen Konstruktionsgrammatik. *Zeitschrift für Sprachwissenschaft* 32 (1): 41–73.
- Halle, Morris & Marantz, Alec (1993) Distributed Morphology and the pieces of inflection. In Kenneth Hale & S. Jay Keyser (eds.), *The View from Building 20*, pp. 111–176. Cambridge: MIT Press.
- Harley, Heidi & Noyer, Rolf (2003) Distributed Morphology. In *The Second GLOT International State-of-the-Article Book*, pp. 463–496. Berlin: Mouton de Gruyter.
- Harnisch, Rüdiger (2003) Verbabhängige Morphologie und Valenzmorphologie der Subjekt-Verb-Kongruenz. In Vilmos Ágel, Ludwig Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft: Dependenz und Valenz*, Vol. 1, pp. 411–421. Berlin: Walter de Gruyter.
- Harris, Zellig (1942) Morpheme alternants in linguistics analysis. *Language* 18: 169–180.
- Hays, David (1964) Dependency theory: A formalism and some observations. *Language* 40: 511–525.
- Heringer, Hans-Jürgen (1970) Einige Ergebnisse und Probleme der Dependenzgrammatik. *Der Deutschunterricht* 4: 42–98.
- (1973) *Theorie der deutschen Syntax*. Second edition. München: Hueber.
- (1996) *Deutsche Syntax Dependentiell*. Tübingen: Staufenberg.
- Hockett, Charles F. (1947) Problems of morphemic analysis. *Language* 23: 321–343.
- (1954) Two models of grammatical description. *Word* 10: 210–234.
- Hudson, Richard (1984) *Word Grammar*. New York: Basil Blackwell.
- (1990) *An English Word Grammar*. Oxford: Basil Blackwell.
- (2003) Word Grammar. In Vilmos Ágel, Ludwig Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft: Dependenz und Valenz*, Vol. 1, pp. 508–526. Berlin: Walter de Gruyter.
- (2007) *Language Networks: The New Word Grammar*. New York: Oxford University Press.
- (2010) *An Introduction to Word Grammar*. Cambridge: Cambridge University Press.
- Jung, Wha-Young (1995) *Syntaktische Relationen im Rahmen der Dependenzgrammatik*. Hamburg: Buske.
- Kahane, Sylvain (ed.) (2000) *Les Grammaires de Dependence*. Paris: Hermes.
- Kunze, Jürgen (1975) *Abhängigkeitsgrammatik*. Berlin: Akademie Verlag.
- Lehmann, Christian (2004) Interlinear morphemic glossing. In Geert Booij, Christian Lehmann, Joachim Mugdan & Stavros Skopeteas (eds.), *Morphologie. Ein internationales Handbuch zur Flexion und Wortbildung. 2. Halbband (Handbücher der Sprach- und Kommunikationswissenschaft, 17.2)*, pp. 1834–1857. Berlin: Walter de Gruyter.

- Lieber, Rochelle (1987) *An Integrated Theory of Autosegmental Processes*. Albany, NY: State University of New York Press.
- Lieber, Rochelle & Scalise, Sergio (2007) The Lexical Integrity Hypothesis in a new theoretical universe. In Geert Booij, Luca Ducceschi, Bernard Fradin, Emiliano Guevara, Angela Ralli & Sergio Scalise (eds.), *On-line Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5)*, pp. 1–24. Bologna: University of Bologna. <<http://www3.lingue.unibo.it/mmm2/wp-content/uploads/2012/09/001-024-Lieber-Scalise.pdf>> (25 August 2014)
- Lobin, Henning (1993) *Koordinationsyntax als prozedurales Phänomen*. Tübingen: Narr.
- Matthews, Peter Hugoe (1972) *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press.
- (1981) *Syntax*. Cambridge: Cambridge University Press.
- (1991) *Morphology: Second Edition*. Cambridge: Cambridge University Press.
- Maxwell, Dan (2003) The concept of dependency in morphology. In Vilmos Ágel, Ludwig Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft: Dependenz und Valenz*, Vol. 1, pp. 678–684. Berlin: Walter de Gruyter.
- McCarthy, John (1981) A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12: 373–418.
- Mel'čuk, Igor (1988) *Dependency Syntax: Theory and Practice*. Albany, NY: State University of New York Press.
- (2003) Levels of dependency in linguistic description: Concepts and problems. In Vilmos Ágel, Ludwig Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans Jürgen Heringer & Henning Lobin (eds.), *Handbücher zur Sprach- und Kommunikationswissenschaft: Dependenz und Valenz*, Vol.1, pp. 188–229. Berlin: Walter de Gruyter.
- Nida, Eugene A. (1948) The identification of morphemes. *Language* 24: 414–441.
- Nivre, Joakim (2006) *Inductive Dependency Parsing*. Berlin/New York: Springer.
- O'Grady, William (1998) The syntax of idioms. *Natural Language and Linguistic Theory* 16: 79–312.
- Osborne, Timothy (2005) Beyond the constituent: A DG analysis of chains. *Folia Linguistica* 39 (3– 4): 251–297.
- Osborne, Timothy & Groß, Thomas (2012a) Constructions are catenae: Construction Grammar meets dependency grammar. *Cognitive Linguistics* 23 (1): 163–214.
- (2012b) Antecedent containment: A dependency grammar solution in terms of catenae. *Studia Linguistica* 66 (2): 94–127.
- Osborne, Timothy; Putnam, Michael & Groß, Thomas (2011) Bare phrase structure, label-less trees, and specifier-less syntax: Is Minimalism becoming a dependency grammar? *The Linguistic Review* 28: 315–364.
- (2012) Catenae: Introducing a novel unit of syntactic analysis. *Syntax* 15 (4): 354–396.
- Palmer, Frank R. (2001) *Mood and Modality: Second Edition*. Cambridge: Cambridge University Press

- Pickering, Martin & Barry, Guy (1993) Dependency categorial grammar and coordination. *Linguistics* 31: 855–902.
- Robins, Robert H. (1959) In defense of WP. *Transactions of the Royal Philological Society*: 116–144.
- Robinson, Jane (1970) Dependency structures and transformational rules. *Language* 46: 259–285.
- Sadock, Jerrold M. (1991) *Autolexical Syntax: A Theory of Parallel Grammatical Relations*. Chicago/London: Chicago University Press.
- Schubert, Klaus (1988) *Metataxis: Contrastive Dependency Syntax for Machine Translation*. Dordrecht: Foris.
- Sgall, Petr; Hajičová, Eva & Panevová, Jarmila (1986) *The Meaning of the Sentence in its Semantic and Pragmatic Aspects*. Dordrecht: D. Reidel.
- Spencer, Andrew (1991) *Morphological Theory*. Oxford: Blackwell.
- Starke, Michal (2009) Nanosyntax: A short primer to a new approach of language. *Nordlyd* 36.1, pp. 1–6. <<http://www.ub.uit.no/baser/nordlyd/>> (26 July 2014)
- Starosta, Stanley (1988) *The Case for Lexicase: An Outline of Lexicase Grammatical Theory*. New York: Pinter Publishers.
- Stump, Gregory T. (2001) *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge: Cambridge University Press.
- Stump, Gregory T. & Finkel, Raphael A. (2013) *Morphological Typology: From Word to Paradigm*. Cambridge: Cambridge University Press.
- Tarvainen, Kalevi (2000) *Einführung in die Dependenzgrammatik*. Tübingen: Niemeyer.
- Tesnière, Lucien (1959) *Éléments de Syntaxe Structurale*. Paris: Klincksieck.
- Williams, Edwin (1981) On the notions ‘lexically related’ and ‘head of a word’. *Linguistic Inquiry* 12: 245–274.
- (2011) *Regimes of Derivation in Syntax and Morphology*. New York: Routledge.

## Abbreviations

Category abbreviations follow the Leipzig Glossing Rules. Additional abbreviations:

- PSS possibility  
 VOL volitional, volitive (following Lehmann 2004)

Contact information:

Thomas Groß  
 Aichi University  
 Nakamura-ku, Hiraike-cho 4-60-6  
 Nagoya, 453-8777 Japan  
 e-mail: [tmgross\(at\)vega\(dot\)aichi-u\(dot\)ac\(dot\)jp](mailto:tmgross@vega.aichi-u.ac.jp)



Wojciech Lewandowski

## Deictic Verbs: Typology, Thinking for Speaking and SLA

### Abstract

This paper investigates the acquisition of the motion verbs COME and GO by Polish speakers of Spanish L2. I show that whereas these verbs encode deictic information in Spanish, in Polish their use relies on non-deictic factors. In particular, COME is preferred when a goal-oriented perspective is adopted, while GO implies an source-oriented conceptualization of the motion event. Following Slobin (1996), I demonstrate that these typological contrasts yield different patterns of thinking for speaking, which influence L2 acquisition even at advanced stages.

### 1. Introduction

This paper is divided into two main parts: on the one hand, it presents a typological analysis of the lexical semantics of the deictic verbs *come* and *go* (henceforth, C&G), focusing mainly on data from Spanish and Polish and, on the other, it investigates on an empirical basis the role of “thinking for speaking” for SLA.

Following Fillmore (1997), I mean by deictic those verbs whose interpretation relies on the spatial and temporal location of the speech act participants. For instance, as shown in (1) the Spanish verbs *venir* ‘come’, *traer* ‘bring’ and *ir* ‘go’, *llevar* ‘take’ are deictic, because the first two imply the presence of the speaker at the goal of movement (which is conveyed by the spatial adverb *aquí* ‘here’), while the last two imply his/her absence at the goal of movement (which is conveyed by the spatial adverb *allí* ‘there’).

- (1) a. *Ven* / *tráe-lo* *aquí* / *\*allí*.  
come.IMP / bring.IMP-it here / there  
‘Come/bring it here/\*there.’

- b. *Ve* / *lléva-lo* *allí* / *\*aquí*.  
 go.IMP / take.IMP-it here / here  
 ‘Go/take it there/\*here.’

Such implications are not involved in the meaning of other verbs, like *entrar* ‘enter’, since their interpretation is independent of the location of the speech act participants.

- (2) *Entr-a* *aquí* / *allí*.  
 enter-IMP here / there  
 ‘Enter here/there.’

Although it has generally been assumed that C&G exhibit a deictic contrast in every language, characterized as motion towards the speaker vs. motion away from the speaker (Miller & Johnson-Laird 1976; Talmy 2000), recent cross-linguistic research has revealed that whereas such an analysis appropriately captures the lexical semantics of C&G in languages such as Spanish or Portuguese, in others, the deictic center of C can be shifted to other goals of movement, e.g. the addressee or even another goal of movement beyond the speech act participants.<sup>1</sup> Moreover, in most Slavic languages, including Polish, the use of C&G is related to other, non-deictic factors (Ricca 1993; Lewandowski 2010).

Drawing on this observation, I show in this article that in Polish (and probably most Slavic languages) C is preferred when the speaker wishes to adopt an arrival-oriented perspective, and G, if the motion event is conceptualized from a source-oriented perspective. As a consequence, Polish speakers can choose to think about the same motion event from two different perspectives (that of arrival or that of departure), while no such possibility is available in languages where C&G codify strict deictic information (e.g., Spanish or Portuguese) and, hence, are in complementary distribution.

---

<sup>1</sup> As suggested by an anonymous reviewer, the Spanish sentence *¿Vendrás mañana a la fiesta de Luis?* (‘Are you coming tomorrow to Luis’ party?’), pronounced in a situation in which neither the speaker nor the hearer are situated at the goal of movement, also illustrate the deictic center shift. Yet, according to my understanding, this so-called comitative context is a particular kind of motion towards the speaker at the reference time (this term will be explained more precisely in Section 2). Concretely, C can be used only if the speaker will be present at Luis’ party (thus, his presence at the goal of movement is implied); otherwise, G has to be used.

In the second part of the article I address the implication of this cross-linguistic divergence for SLA. I adopt Slobin's (1996) "thinking for speaking" hypothesis, according to which different language patterns yield different patterns of thought in the process of expressing and interpreting verbal expressions. Consistently, we hypothesize that Polish speakers of Spanish L2 use C when the speaker strongly identifies with the goal of movement and G, when no such identification takes place, thus violating the strict deictic conditions of use of C&G in this language. By contrast, it can be expected that Spanish speakers of Polish L2 will follow the deictic conditions of use of C&G, not using C in contexts of motion towards a discourse entity different from the speaker, even if the contextual information indicates that the event is conceptualized as arrival-oriented.

The aim of this article is to test the first part of this hypothesis, i.e. the one concerning the acquisition of deictic verbs by Polish speakers of Spanish L2. The hypothesis was tested by means of an acceptability judgment task carried out with 30 Polish learners of Spanish and the results were interpreted in the light of Slobin's (1996) "thinking for speaking" hypothesis.

This article is organized as follows: Section 2 focuses on the usage patterns of C&G in Polish and Spanish. Section 3 and Section 4 investigate the influence of the typological differences reported in Section 2 for SLA. In particular, after giving a brief overview in Section 3 of the theoretical background of my study, that is, the "thinking for speaking" hypothesis and its application to SLA, I formulate in Section 4 the hypothesis, describe the procedure of the experiment and report on the results. Conclusions are drawn in section 5.

## **2. C&G in Polish and Spanish: Different ways of spatial conceptualization**

As already noted in the Introduction, it has been widely assumed in the literature on motion events that all languages have a class of motion verbs corresponding to English *come* and *go* and that these verbs display a universal deictic contrast (cf. Miller & Johnson-Laird 1976; Talmy 2000, among many others).

However, as has been argued by Lewandowski (2010), there exist important cross-linguistic differences in the lexical semantics of C&G. While it is true that in some languages C describes motion towards the speaker, in others, the deictic center can be extended to other goals of

movement, such as the addressee or even another goal of movement beyond the speech act participants (Gathercole 1977; Ricca 1993; Di Meola 1994, among others). Quite importantly, the range of the possible goals of movement which can be codified as the Ground in C, seems to be subordinated to a strict universal hierarchy, which is shown in Table 1.

**Table 1.** Hierarchy of Grounds lexicalized in C

Goal	Languages
1. the speaker's location at the coding time	Portuguese, Shibe, ...
2. the speaker's location at the reference time	Jacatlec, Spanish, ...
3. the addressee's location	Catalan, English, Nepali, Turkish, ...
4. another goal of movement	Czech, Polish, Russian, ...

Following Fillmore (1971: 52), the term “coding time” is used to refer to the time of the communication act, and “reference time” to describe the time of the spatial event. Thus, in (3a) the spatial adverb “here” implies that the utterance denotes motion toward the place where the speaker is located at coding time, i.e., the moment of speaking; by contrast, in (3b) motion towards the speaker's location at reference time is referred to, since the deictic pronoun “that” conveys that he/she is located in a different place when uttering the sentence.

- (3) a. *John came here yesterday.*  
 b. *John came to visit me at that place.*

Turning now to the semantic hierarchy depicted in Table 1, Lewandowski (2010) shows that C which can take as the Ground a goal of movement situated lower in the hierarchy than the speaker's location at the coding time, automatically allows for any other goal, which is placed higher in the established hierarchy. For example, in Portuguese, C is allowed only when motion toward the speaker is referred to at the coding time. In Spanish, C is used both in contexts of motion towards the speaker at the coding time and at the reference time. In Catalan, the verb under discussion describes displacement toward the speaker (at the coding and reference time) and the hearer. Finally, in languages such as Polish, C can refer to movement toward ANY goal (the speaker, the hearer, or a place situated beyond the speech act participants).

On the other hand, cross-linguistic data analyzed by Ricca (1993) strongly suggests that, usually, if such an extension of the deictic center takes place, C&G tend to alternate. For the sake of clarity, this issue can be

illustrated with some examples from English. All of them are taken or adapted from Fillmore (1997).<sup>2</sup>

- (4) a. *He came/\*went here two hours before I arrived.* (Goal 1)  
 b. *He'll come/go to the office tomorrow to pick me up.* (Goal 2)  
 c. *She'll come/go there to meet you.* (Goal 3)  
 d. *Tomorrow, I'll go/\*come to John's place.* (Goal 4)

Since in English the deictic center can be shifted from the speaker's location at coding time (4a) to the addressee's location, both verbs alternate in the deictic center extension zone, i.e. in contexts of motion toward the speaker's location at reference time (4b) and motion toward the addressee (4c), but not in the context of motion toward another goal (4d). As shown in Table 1, in Spanish, such deictic center extension zone is constituted by Goal 2. As will be shown, under this condition of use of C&G, the alternation between both verbs is also possible, although G is clearly preferred when the absence of the speaker at the goal of movement is implied. Moreover, Fillmore (1997) noted that C&G differ not only as to the deictic information, but they also codify a different type of temporal orientation: G is source-oriented, since the temporal specification in (5a) refers to the initial point of movement, whereas C is goal-oriented, since the temporal specification in (5b) refers to the arrival time.

- (5) a. *I went home at seven.*  
 b. *I came home at seven.*

One crucial conclusion should be drawn at this point: while the deictic center shift possible in some languages allows the adoption of two different perspectives (or construals, in Langacker's (1987) terms) when referring to the same objective spatial situation (the perspective of departure or the perspective of arrival), such a possibility is not available in languages where C&G codify strict deictic information and thus are in complementary distribution. This phenomenon is clearly related to Slobin's (1996) idea that the semantic and grammatical resources of a particular language influence the way the speaker can choose to think about a given event or entity. In order to illustrate more concretely this semantic divergence, in what follows I provide a more exhaustive comparison of the

---

<sup>2</sup> For further cross-linguistic evidence, see Ricca (1993).

lexical semantics of C&G in Polish and Spanish. In particular, I show that these languages represent two typologically different usage patterns of C&G: whereas in the latter these verbs express the deictic opposition “motion towards the speaker” vs. “motion away from the speaker”, in the former, the use of one or other verb relies on pragmatic factors related to a particular kind of conceptualization of the motion event.

## 2.1 C&G in Spanish

As illustrated in (6a) and (6b), the Spanish verb *venir* ‘to come’ typically describes motion towards the speaker’s location at either coding or reference time, whereas the verb *ir* ‘to go’ is used in the context of movement toward any other goal.

- (6) a. *Ven* / \**ve* *aquí a las cuatro.*  
 come.IMP / go.IMP here at ART four  
 ‘Come/\*go here at four.’
- b. *¿Quién vendr-á / ir-á a ver-nos a ese lugar*  
 who come-3SG.FUT / go-3SG.FUT to see.INF-PRON.1PL to that place  
*tan lejano?*  
 so far-off  
 ‘Who will visit us in that far-off place?’

The spatial adverb *aquí* (‘here’) in (6a) indicates that the speaker is present at the goal of movement at the time when the sentence is uttered (“coding time” in Fillmore’s (1971: 52) terms). Yet, the example in (6b) demonstrates that *venir* can describe not only motion toward where the speaker is located when uttering the sentence, but also toward the speaker’s location at the time of the displacement (“reference time” in Fillmore’s (1971: 52) terms), that is, toward a place where the speaker will be located when the displacement takes place.

The Spanish verb *ir* ‘to go’ is in almost complementary distribution with *venir* ‘to come’ since, as illustrated in (6a) and (6b), it cannot describe scenes where the Figure moves toward the speaker’s location, unless motion toward the speaker’s location at the reference time is denoted. This is in keeping with the typological introduction to deictic verbs outlined above: since the deictic center can be extended in Spanish to Goal 2 in Table 1, under this condition the use of both C as well as G is allowed.

However, it should be pointed out that the use of one or other verb involves a very important difference in meaning. As shown in (7a), G typically implies the speaker's absence at the goal of movement: since in (7a) the *Pluscuamperfecto* (Past Perfect) conveys that the speaker was not located at the goal of movement (the library) at the time when his/her brother arrived there, native speakers tend to use G. By contrast, (7b) describes a scene where the speaker is certainly present at the goal of movement (the airport) at the reference time and, consequently, G sounds odd in this sentence.

- (7) a. *Llegu-é a la biblioteca y vi que también*  
 arrive-1SG.PST to the library and see.1SG.PST that also  
*hab-ía ido /<sup>??</sup>venido mi hermano.*  
 AUX-3SG.PST gone / come my brother  
 'When I arrived at the library, I realized that my brother had gone/<sup>??</sup>come there, too.'<sup>3</sup>
- b. *He telefonado desde el aeropuerto y me han dicho*  
 AUX.1SG called from the airport and me AUX.3PL told  
*que ven-ían /<sup>??</sup>i-ban a buscar-me.*  
 that come-3PL.PST / go-3PL.PST to pick up.INF-PRON.1SG  
 'I called from the airport and they told me that they were coming/<sup>??</sup>going to pick me up.'

A special case of motion towards the speaker are the so-called comitative contexts, i.e. situations in which the speaker asks the addressee to accompany him/her to a place. In Spanish, the use of C is obligatory in such speech acts (see (8)), since they involve, first of all, the addressee's displacement towards the speaker, while the displacement of both to another goal of movement may be regarded as a less salient property of their illocutionary force (the term is used in the sense of Austin 1975 and Holdcroft 1978).

- (8) *¿Te vien-es con-migo a-l cine?*  
 PRON.2SG come-2SG.PRES with-me to-the cinema  
 'Would you like to come with me to the cinema?'

---

<sup>3</sup> I am grateful to José María Brucart for the examples in (7).

To sum up, in Spanish the verb *venir* describes purely motion towards the speaker (at either the coding or the reference time), whereas *ir* refers to motion in a direction different from the speaker.

## 2.2 C&G in Polish

The system of the Polish C&G is more complex than in Spanish since at least two Polish verbs corresponding to the Spanish *venir* can be found in dictionaries: *przyjść*, denoting movement on foot and *przyjechać*, denoting movement by vehicle (cf. Las & Wasilenko 2006). The same applies to the equivalents of *ir*: *pójść* refers to motion on foot, whereas *pojechać* refers to motion by vehicle. For the sake of simplicity, all examples cited in this subsection contain the verbs referring to motion on foot, but the same conditions of use are valid for the verbs referring to motion by vehicle.

As shown in (9),<sup>4</sup> both types of verbs can be used in contexts of motion towards any goal, i.e. the speaker (9a, 9b), the addressee (9c, 9d) or a goal beyond the speech act participants (9e) independently of whether the speaker is or is not located at the goal of movement at the coding or reference time.

- (9) a. *Jan przyszedł wczoraj do mnie.*  
 John come.3SG.PST yesterday to me.GEN  
 ‘John came to my place yesterday.’
- b. *Powiedział jej, że byłem chory i żeby do mnie poszła.*  
 tell.3SG.PST her that be.1SG.PST ill and that to me.GEN go-3SG.PST  
 ‘He told her that I was ill and asked her to come to my place.’
- c. *Mówi-leś, że jak ktoś do ciebie*  
 say-2SG.PST that when somebody.NOM to you.GEN  
*przychodzi, robi-sz się nerwowi.*  
 come.3SG.PRES make-2SG.PRES REFL nervous  
 ‘You said that every time somebody goes to your place, you get nervous.’

---

<sup>4</sup> The data illustrating the use of C&G in Polish in (9) are introspective.

- d. *Najpierw posz-li do biura i stamtąd posz-li*  
 first go-3PL.PST to office.GEN and from there go-3PL.PST  
*do ciebie.*  
 to you.GEN  
 ‘First they went to the office and from there they went to your place’
- e. *Podoba-ła Ci się impreza w Krakow-ie?*  
 like-3SG.PST you.DAT REFL party.NOM in Cracow-LOC  
*Przysz-ło dużo ludzi?*  
 come-3SG.PST many people.GEN  
 ‘Did you enjoy the party in Cracow? Did many people go there?’

As a general rule, C is preferred when the speaker strongly identifies with the goal of movement, and G when the speaker focuses on the departure point. As shown in (9a), usually when motion towards the speaker is described in a neutral context, the use of C is preferred: since the speaker is at the goal of motion, it is natural for him/her to take his/her own – arrival-oriented – perspective. However, the use of G is possible, e.g., when the speaker wishes to convey that he or she identifies with the source-oriented perspective of the person whose words he/she reports, as in (9b).

In (9c) and (9d) the goal of motion is constituted by the addressee. In (9c) the arrival-oriented perspective is taken, because, as in (9b), the speaker relates the event from the viewpoint of the subject of the sentence, i.e. the person whose message is described. However, in (9d) the departure perspective is due to the source-Path expression “from there”, which determines the spatial orientation of the utterance.

In (9e), the speaker is talking about a party at a place he did not go to, but he uses C, because the goal of movement has previously been introduced in the discourse and so it serves as a focal Ground of the narration in the mind of the speaker.<sup>5</sup>

And finally, let us recall that in contrast to Spanish, comitative contexts in Polish require the adoption of a departure perspective, since, as illustrated in (10), in such speech acts the use of G is obligatory.

- (10) *Pójdiesz ze mną do kin-a?*  
 go.2SG.FUT with me.DAT to cinema-GEN  
 ‘Would you like to come with me to the cinema?’

<sup>5</sup> It is important to stress that in Spanish (9c) and (9e) are possible only if the speaker’s presence at the goal of movement is implied at the coding or reference time, while such conditions are not required in Polish.

Summarizing this section, in Polish, it is possible to adopt two different perspectives (or construals, in Langacker's (1987) terms) when referring to the same objective spatial situation, that is, the perspective of departure or the perspective of arrival. No such possibility is available in the case of the Spanish C&G where motion towards the speaker can be depicted solely from the perspective of the arrival point (C is obligatory), whereas motion towards any other goal must be described from the perspective of the departure point (G is required). Let us recall that this phenomenon clearly reflects Slobin's (1996) idea that the resources of a given language determine (to a certain extent) the way the speaker can choose to think about a particular event when speaking about it.

The remainder of this paper explores the implication of these typologically divergent usage patterns for SLA. In particular, I will report on an acceptability judgment task related to the acquisition of C&G by Polish learners of Spanish. However, before going into details of my experimental study, a brief overview of the theoretical background of my experiment needs to be provided.

### 3. Thinking for speaking in SLA

Over the last thirty years, many studies have explored the coding of motion expressions from a cross-linguistic perspective. As is well known, Talmy (1975, 1985, 2000) first proposed a typology of verbal lexicalization patterns according to which languages can be categorized as either verb-framed, such as Spanish, Turkish, Basque, etc., or satellite-framed, such as English, German, Polish, etc. In the former, the verbs express the Path component typically in the verb stem, while the latter commonly encode in the verb stem the Manner, with the Path being relegated to a secondary element, commonly a preposition or a prefix (cf. (11)).

- (11) a. *The bottle floated into the cave.* (English)  
 b. *La botella entró a la cueva flotando.* (Spanish)  
 the bottle entered to the cave floating  
 (Talmy 1985: 69ff)

On the other hand, research on first language acquisition has demonstrated that children learning typologically different languages pay attention to different aspects of motion events when talking about them, revealing that the influence of the linguistic input the children are exposed to is quite

strong as far as perception and expression of spatial relations is concerned (e.g. Berman & Slobin 1994; Bowerman 1994, 1996; Slobin 1996 *contra* Piaget & Inhelder 1956; Levine & Carey 1982, which focused on the role of the nonlinguistic experience in the development of spatial language). For instance, Slobin (1996) found that, when speaking about motion, English-speaking children used twice as many manner verbs as Spanish-speaking children. They also provided richer descriptions of paths. However, Spanish native speakers, when talking about the same motion event, tended to pay more attention to aspects of the static scene in which the movement took place. This has been attributed to the fact that speakers of satellite-framed languages pay more attention to the conflation of Motion and Manner than speakers of verb-framed languages in which Motion is usually conflated with the Path component in spatial expressions.<sup>6</sup>

In order to account for this finding, Slobin (1996) coined the term “thinking for speaking”, which is defined as “a special form of thought that is mobilized for communication” (Slobin 1996: 76). Importantly, the “thinking for speaking” hypothesis differs crucially from the controversial linguistic determinism hypothesis (Whorf 1956): whereas the latter states that different language patterns yield different patterns of thought (i.e. the language a person speaks shapes the way in which this person understands the world), the former refers only to the role of language in the process of expressing and interpreting verbal expressions without predicting anything about the influence of language in thought in general.

As rightly observed by Stam (1998), if it is true that linguistic categories play an important role in the shaping of concepts that children are going to use in speaking, this would mean that learning a typologically different L2 involves learning another pattern of “thinking for speaking”. Most recently, Cadierno and Lund (2004) suggested that applying the “thinking for speaking” theory for SLA would allow for a systematic investigation of the role of the learners’ L1 in their acquisition of (motion events in) L2.

Researchers came to apparently contradictory conclusions concerning this issue. Cadierno and Ruiz (2006) examined the expression of Path and Manner of motion by two sets of learners: Danish learners of Spanish, i.e. learners whose L1 and L2 belong to different typological patterns (satellite-

---

<sup>6</sup> As suggested by an anonymous reviewer, the reason behind the mentioned attention bias could be that satellite-framed languages have more verbs expressing manner at their disposal than verb-framed languages (cf. Kopecka 2006).

framed and verb-framed) and Italian learners of Spanish, i.e. learners whose L1 and L2 belong to the same typological pattern (the verb-framed one). The findings of the analysis supported the view of a rather limited role of L1 “thinking for speaking” patterns in advanced L2 Spanish learners. In fact, more significant differences were found between the Spanish L1 group and the two groups of learners than between the two learner groups themselves: contrary to the predictions, the Danish learner group did not use alternative means of expressing manner information to a greater extent than the Italian learner group (although it did exhibit a higher degree of elaboration of the semantic component of Path of motion than the Italian and the Spanish NS groups by incorporating a larger number of Ground specifications). However, in an independent line of research, Carroll, von Stutterheim and Nüse (2004) showed that L1 predispositions for “thinking for speaking” persist even at advanced stages. They compared the temporal framing of events by German speakers of English and found that even advanced learners tended to use the German narration pattern, consisting of sequences of bounded events of the type “he walks and then he sees... and then he thinks...”, etc., whereas English native speakers presented the events from an ongoing perspective with respect to the deictic reference point “now” (“he is walking... and he sees... and then he is thinking...”, etc.), which is due probably to the fact that German has no grammaticized progressive aspect.

The remainder of the present paper aims to provide evidence from the acquisition of C&G by Polish learners of Spanish supporting the idea that there is a L1 “thinking for speaking” influence in SLA, as Carroll, von Stutterheim and Nüse (2004) claim, but that certain patterns are acquired more easily than others. In other words, I want to argue that the relevant question is not *whether* there is a L1 “thinking for speaking” influence in SLA, but rather *why* certain “thinking for speaking” patterns are more difficult to restructure than others.

## **4. The present study**

### **4.1 Hypothesis**

Taking into account the considerable cross-linguistic differences in the use patterns of C&G in Polish and Spanish, I hypothesized that there would be a L1 “thinking for speaking” influence in the acquisition of deictic verbs, i.e. that Polish learners of Spanish would use C in situational contexts

focusing on the goal of movement, and G in situational contexts focusing on the source of movement. This hypothesis involves five different situations, two in which a correct use (Conditions 1 and 4) and three in which an incorrect use of C&G is expected (Conditions 2, 3 and 5). I summarize these conditions in Table 2.

**Table 2.** Conditions of use of C&G in Polish and Spanish

GOAL OF MOTION	POLISH	SPANISH
1. SPEAKER, neutral context	C	C
2. SPEAKER, departure perspective	G	C
3. SPEAKER, comitative context	G	C
4. NON-SPEAKER, departure perspective	G	G
5. NON-SPEAKER, arrival perspective	C	G

## 4.2 Participants

The experimental group consisted of 40 learners of Spanish at two Polish universities: the Jagiellonian University in Cracow and the Adam Mickiewicz University in Poznań. 11 of them were male, 29 female and they were between 19 and 25 years old. For a participant's data to be included, all the following criteria had to be met: (a) their level of Spanish had to be "low intermediate" in order to ensure that they could understand the experimental items (see 4.3. "Testing instruments"), (b) they had to be native speakers of Polish, (c) the participant had not have stayed in a bilingual Spanish autonomous region for more than ten days, and (d) the participant had not to know any other Romance language. The last two criteria were included to prevent possible interference with other Romance or peninsular languages, such as Catalan or French, in which C&G behave differently than in Spanish. Ten native speakers of Spanish acted as controls: All of them were living in Spain at the time of the experiment. None of them were bilingual, Spanish was their only reported native language and they had never lived in a bilingual Spanish autonomous region.

## 4.3 Testing instruments

The testing instruments consisted of: (i) a language experience questionnaire, (ii) a cloze test testing their level of Spanish (see Appendix 1) and (iii) an acceptability judgment task designed to test the hypothesis stated in 4.1 (see Appendix 2).

The aim of the language experience questionnaire was to exclude from the analysis those participants which did not meet the criteria given in 4.2. Of the original 40 subjects, 1 reported that his/her native language was not Polish, 5 reported that they were learning another Romance language and 2 had stayed for more than 10 days in Catalonia. These subjects were not included in further stages of the experiment.

I elaborated my own cloze test in order to group the subjects of the experiment into different levels of language knowledge. It consisted of 100 items based on the lexical and grammatical material available in the “Sueña” books for learning Spanish as a foreign language corresponding to levels A1 and A2 (Álvarez Martínez, Blanco Canales, Gómez Sacristán, et al. 2001), B1 (Cabrerizo Ruiz, Gómez Sacristán & Ruiz Martínez 2006), B2 (Álvarez Martínez, De la Fuente Martínez, Giraldo Silveiro, et al. 2007) and C1 (Blanco Canales, Fernández López & Torrens Álvarez 2007). Each level was represented in the test by 20 items (10 grammatical and 10 lexical). I considered participants with more than 40 correct responses as “low intermediate”, with more than 60 correct responses as “high intermediate” and with more than 80 correct responses as “advanced”. Table 3 summarizes the means, standard deviation and ranges for the cloze test together with the mean age and the number of subjects within each level of language knowledge.

**Table 3.** Results of the cloze test

Level	M Results	SD	Range	M age	Number of subjects
Advanced	88.6	8.105	81-100	23.5	8
High Intermediate	73.6	4.776	61-80	22.4	12
Low Intermediate	48.2	5.788	41-60	20.4	10
Other	18.5	4.949	0-40	19	2

As seen from the results of the cloze test, the level of 2 of the remaining 32 participants was below intermediate. Their results were excluded from the analysis.

And finally, the main hypothesis was investigated using an acceptability judgment task, consisting originally of 16 sentences, 4 items per Condition (2 correct and 2 incorrect sentences), and the same number of distractors. The experimental items were intended to strongly reflect the five conditions depicted in Table 2. However, for methodological reasons I decided in the acceptability judgment task to test the preference of Polish native speakers for one or other verb in each context independently by means of a translation into Polish. This was performed with a control group

consisting of 10 subjects. All of the participants gave the expected responses for Conditions 1, 3, 4 and 5. But, 4 out of 10 Polish native speakers preferred the use of C to G under the second condition, which was removed from the experimental task.

The vocabulary used in each sentence fitted the lexical minima established in “Sueña” for low intermediate learners of Spanish in order to avoid the possibility that learners of low proficiency levels rejected sentences because of a word unknown to them or judged the sentence wrong because of a wrong understanding of the context. Given the fact that the order of the sentences might influence the result of the experiment owing to factors such as nervousness at the beginning of the experiment or fatigue toward the end of the experiment (Schütze 1996), the order of the experimental items was counterbalanced across different participants. Below each sentence there was a Likert scale with values from -2 to +2 in order to judge a given item as “sounds awkward” (-2), “sounds bad” (-1), “I don’t know” (0), “sounds okay” (+1) and “sounds perfect” (+2). I considered a Likert scale questionnaire a convenient data collection instrument as it is practical to use with large numbers of participants and offers clear numerical data which are easy to analyze. Afterwards, the original Likert scale was converted to binominal data by combining all answers into the two categories of “correct” and “incorrect”. For example, when the experimental item was a correct sentence and it was judged as either “sounds okay” or “sounds perfect”, both answers were considered “correct”. By contrast, when such a sentence was judged as “sounds awkward” or “sounds bad”, both answers were assigned the value “incorrect”.

#### **4.4 Procedure**

The data were collected between 3rd of August and 5th of September 2009 via an on-line video conference. First, all the participants answered the language experience questionnaire and after that they took the cloze test, measuring their level of Spanish. Next, they were presented with the acceptability judgment task and asked to judge the sentences on a Likert scale from -2 for “completely unacceptable” to +2 for “perfectly acceptable”, according to their first impression. Precise instructions with examples not related to the sentences at issue were provided in order to explain the reasons why a sentence should be considered acceptable or unacceptable. The participants took 15-20 minutes to complete the

acceptability judgment task and they were not allowed to go back and modify their responses.

#### 4.5 Results and discussion

Since it is controversial whether parametric tests such as ANOVA are appropriate for measuring Likert-scale data, the non-parametric Kruskal-Wallis and Fisher Exact Probability tests were run in order to analyze the data. The statistical software R was employed for this purpose.

As a first step, the Kruskal-Wallis test was run in order to investigate whether some of the conditions of use were acquired correctly by all experimental groups. According to the results there are no statistically significant differences between groups for Conditions 1 and 4 ( $p=0.3916$  with 3 df for C1, while  $p=0.1646$  with 3 df for C4). This was expected, since in these situations the conditions of use of C&G in Polish overlap with the Spanish ones, from which it logically follows that the “thinking for speaking” patterns did not need to be restructured: Polish speakers of Spanish judged correctly the experimental items from the earliest stages of acquisition because the “thinking for speaking” patterns are identical in both languages.

By contrast, for Conditions 3 and 5, the overall between-groups effect is significant ( $p>0.05$  in each case), which means that there is at least one proficiency level whose responses differ in statistical terms from the responses of the control group. Thus, as a second step, a multiple comparison between groups for Conditions 3 and 5 was carried out by means of the Fisher Exact test. The Pearson Chi-Squared test could not be used due to the fact that the size of the samples was too small. With regard to Condition 5, the results show that it is practically impossible for Polish native speakers to learn this condition of use, since even advanced learners judge the experimental items incorrectly. The Fisher Exact test reveals, on the one hand, that the answers of each of the experimental groups differ significantly from the answers of the control group ( $p>0.01$  for all cases). On the other hand, there are no significant differences between G1, G2 and G3, which means that no progress in the acquisition of the condition under discussion was made. The p-values for the between-groups comparison are shown in Table 4.

**Table 4.** The between-groups comparison for Condition 5

G1-G2	G1-G3	G1-G4	G2-G3	G2-G4	G3-G4
p=0.1028	p=1	p<0.01	p=0.2542	p<0.01	p<0.01

In contrast to Condition 5, clear progress in the acquisition of Condition 3 was observed. Although there is a significant difference between the answers of the control group vs. G1 and G2 (low and high intermediate), the answers of the advanced group (G3) and the control group are equal from a statistical viewpoint, though the effect is not very robust ( $p=0.0471$ ). On the other hand, one should be aware that there are significant differences between G1 and G3 as far as the correctness of the judgments is concerned (see Table 5), which means that correct answers increase significantly with the proficiency level.

**Table 5.** The between-groups comparison for Condition 3

G1-G2	G1-G3	G1-G4	G2-G3	G2-G4	G3-G4
p=0.111	p<0.01	p<0.01	p=0.3203	p<0.01	p=0.0471

Concluding, my hypothesis is partially borne out: undoubtedly, there is a L1 “thinking for speaking” influence in the acquisition of C&G by Polish speakers of Spanish L2 even at advanced stages. However, some mappings of formal expression and conceptual content are easier to restructure than others (cf. Condition 5 vs. Condition 3). Although more research is necessary in order to answer the question of why Condition 3 is more prone to be acquired by Polish learners of Spanish than Condition 5, I would like to suggest that this might be due to the goal-bias in human cognition. In particular, it has been shown that children and adults tend to encode Goal paths in preference to Source paths and that asymmetry in speech production could have its origins in non-linguistic event representations (Lakusta & Landau 2005; Lakusta, Wagner, O’Hearn, et al. 2007). As far as the results of my experiment are concerned, it is important to note that Condition 3 involves the shift from a source-oriented perspective to a goal-oriented perspective (in Polish the departure-oriented G is used, whereas in Spanish the arrival-oriented C is obligatory). By contrast, Condition 5 involves the shift from a goal-oriented perspective to a source-oriented perspective (in Polish C is preferred, while in Spanish the use of G is required). Thus, it appears that the acquisition of C&G is constrained by the source-goal asymmetry in human cognition, since my experiment

showed that it is possible for Polish learners of Spanish to adopt a new “thinking for speaking” pattern focusing on the final point of movement, but it is almost impossible for them to restructure a goal-oriented perspective and adopt a source-oriented one.

When comparing my study with Cadierno and Ruiz’s (2006) and Carroll, von Stutterheim and Nüse’s (2004), it becomes clear that L1 “thinking for speaking” plays a crucial role in the acquisition of a L2, but it is constrained by different types of factors, such as probably the goal-bias in the case of C&G. However, further research is clearly needed in order to shed light on the nature of such factors and their systematic influence in the acquisition of “thinking for speaking” in L2. In short, the relevant question is not *whether* there is a L1 “thinking for speaking” influence in SLA, but rather *why* certain “thinking for speaking” patterns are easier or more difficult to restructure than others.

Finally, it should be mentioned that according to the anonymous reviewers of the present paper it is a controversial fact whether the “thinking for speaking” theory can be used to explain the results of an acceptability judgment task. As is well known, a typical “thinking for speaking” experiment consists of elicited narratives, i.e., two groups of participants from two different language groups look at the same external reality and describe it in different ways, devoting more attention to some details of the scene and less to others. Although my experiment does not start from a common external stimulus which is further conceptualized linguistically by speakers of two typologically different languages, I think that it is possible to reason that “thinking for speaking” plays a role here. As Slobin’s theory states, language and thought are separate, because there is a lot of thinking which we cannot express even in our mother tongue. However, when we decide to encode our thinking in language, we are forced to shape it according to the grammatical and lexical devices of this language. In the light of this theory it could be argued that it is probably because of the L1 “thinking for speaking” patterns (together with the Goal-bias in human cognition) that native speakers of Polish are practically not able to linguistically conceptualize motion towards a person beyond the speech act participants in an arrival-oriented context by means of G. Although my findings seem to be confirmed in spontaneous speech, where even very advanced Polish speakers of Spanish use C in the context I have already mentioned (indeed, these observations served as an inspiration for the present study), an experiment consisting of elicited narratives is without

any doubt necessary in order to further verify in a statistically measurable fashion the hypothesis of the present paper.

## 5. Conclusions

In this paper, I analyzed the semantics of C&G from a typological perspective. Drawing on Lewandowski's (2010) and Ricca's (1993) work, I showed that important cross-linguistic differences are implied in the meaning of these verbs (*contra* Miller & Johnson-Laird 1976 and Talmy 2000, *inter alia*): whereas in some languages, such as Spanish or Portuguese, they codify strict deictic information concerning the spatial position of the speaker ("motion towards the speaker" vs. "motion away from the speaker"), in others the deictic center of C can be shifted to other goals of movement, e.g. the addressee or even a goal of movement beyond the speech act participants. Quite importantly, in most Slavic languages, including Polish, the use of C&G does not rely on non-deictic factors. In particular, in Polish C is preferred when the speaker wishes to adopt an arrival-oriented perspective and G, if the motion event is conceptualized from a source-oriented perspective. As a consequence, Polish speakers can choose to think about the same motion event from two different perspectives (that of arrival or that of departure), while no such possibility is available in languages where C&G codify strict deictic information.

In the second part of the article I addressed the implications of this cross-linguistic divergence for SLA. I adopted Slobin's (1996) thinking for speaking hypothesis, according to which different language patterns yield different patterns of thought in the process of expressing and interpreting verbal expressions. The results of the experiment led me to conclude that the role of L1 thinking for speaking in L2 is important even at advanced stages, but it is constrained by some kind of non-linguistic factors, which should be investigated with more detail in the future. Probably, one important factor constraining the acquisition of C&G is the source-goal asymmetry in human cognition.

Finally, as far as future research is concerned, this should without any doubt be directed toward (i) learning and teaching deictic verbs, since, as has been shown, the correct use of C&G in some very specific contexts is almost impossible to acquire; (ii) productive use of language (e.g., elicitation of narratives), which would highlight possible contextual differences in the correct-incorrect use of deictic verbs of L2 learners, and

(iii) the relation between deictic verbs, SLA and source-goal asymmetry in human cognition.

## References

- Álvarez Martínez, María Ángeles; Blanco Canales, Ana; Gómez Sacristán, María Luisa & Pérez De la Cruz, Nuria (2001) *Sueña 1*. Madrid: Grupo Anaya.
- Álvarez Martínez, María Ángeles; Vega De la Fuente Martínez, María; Giraldo Silverio, Inocencio; Martín Martín, Fátima; Sanz Sánchez, Begoña & Torrens Álvarez, María Jesús (2007) *Sueña 3*. Madrid: Grupo Anaya.
- Austin, John (1975) *How to Do Things with Words*. Oxford: Oxford University Press.
- Berman, Ruth & Slobin, Dan (1994) *Relating Events in Narrative: A Crosslinguistic Development Study*. Mahwah, NJ: Lawrence Erlbaum.
- Blanco Canales, Ana; Fernández López, María Carmen & Torrens Álvarez, María Jesús (2007) *Sueña 4*. Madrid: Grupo Anaya.
- Bowerman, Melissa (1994) From universal to language-specific in early grammatical development. *Philosophical Transactions of the Royal Society of London B* 346: 37–45.
- (1996) The origins of childrens' spatial semantic categories: cognitive versus linguistic determinants. In John Gumperz & Steven Levinson (eds.), *Rethinking Linguistic Relativity*, pp. 145–176. Cambridge: Cambridge University Press.
- Cabrerizo Ruiz, María Aránzazu; Gómez Sacristán, María Luisa & Ruiz Martínez, Ana María (2006) *Sueña 2*. Madrid: Grupo Anaya.
- Cadierno, Teresa & Lund, Karen (2004) Cognitive linguistics and second language acquisition: motion events in a Typological Framework. In Bill Van Patten, Jessica Williams, Susanne Rott & Mark Overstreet (eds.), *Form-meaning Connections in Second Language Acquisition*, pp. 139–154. Hillsdale: Lawrence Erlbaum.
- Cadierno, Teresa & Ruiz, Lucas (2006) Motion events in Spanish L2 acquisition. *Annual Review of Cognitive Linguistics* 4: 183–216.
- Carroll, Mary; von Stutterheim, Christiane & Nüse, Ralf (2004) The language and thought debate: a psycholinguistic approach. In Thomas Pechmann & Christopher Habel (eds.), *Multidisciplinary Approaches to Language Production*, pp. 183–218. Berlin: Mouton de Gruyter.
- Di Meola, Claudio (1994) *Kommen und gehen. Eine kognitiv-linguistische Untersuchung der Polysemie deiktischer Bewegungsverben* [Come and Go. A Cognitive Linguistics Study of the Polysemy of Deictic Motion Verbs]. Tübingen: Max Niemeyer Verlag.
- Fillmore, Charles (1971) *Santa Cruz Lectures on Deixis*. Bloomington: Indiana University.
- (1997) *Lectures on Deixis*. Stanford: CSLI.
- Gathercole, Virginia (1977) A study of coming and going verbs of the speakers of four languages: Spanish, Japanese, English and Turkish. *Kansas Working Papers in Linguistics* 2: 61–94.

- Holdcroft, David (1978) *Words and Deeds. Problems in the Theory of Speech Acts*. Oxford: Oxford University Press.
- Kopecka, Anetta (2006) Lexicalization of manner of motion: typology and lexical diversity. Paper presented at *AELCO*, October, 2006.
- Lakusta, Laura & Landau, Barbara (2005) Starting at the end: The importance of goals in spatial language. *Cognition* 96: 1–33.
- Lakusta, Laura; Wagner, Laura; O’Hearn, Kirsten & Landau, Barbara (2007) Conceptual foundations of spatial language: evidence for a goal bias in infants. *Language Learning and Development* 3, 3: 179–197.
- Langacker, Ronald (1987) *Foundations of Cognitive Grammar. Theoretical Prerequisites*. Stanford: Stanford University Press.
- Las, Maria & Wasilenko, Magdalena (2006) *Diccionario polaco-español, español-polaco con ejemplos* [A Polish-Spanish and Spanish-Polish Dictionary with Examples]. Berlin/Munich: Langenscheidt Publishers.
- Levine, Susan & Carey, Susan (1982) Up front: the acquisition of a concept and a word. *Journal of Child Language* 9: 645–657.
- Lewandowski, Wojciech (2010) Questioning the universality of deictic oppositions: *come* and *go* in Polish, Spanish and other languages. In Ignasi Navarro & Antonio José Silvestre (eds.), *Language Systems and Cognitive Perspective*, pp. 75–92. Valencia: Tirant lo Blanch.
- Miller, George & Johnson-Laird, Philip (1976) *Language and Perception*. Cambridge: Cambridge University Press.
- Piaget, Jean & Inhelder, Bärbel (1956) *The Child’s Conception of Space*. London: Routledge.
- Ricca, Davide (1993) *I verbi deittici di movimento in Europa: una ricerca interlinguistica* [Deictic Motion Verbs in Europe: A Cross-Linguistic Study]. Firenze: La Nuova Italia Editrice.
- Schütze, Carson (1996) *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Chicago: University of Chicago Press.
- Slobin, Dan (1996) From ‘thought and language’ to ‘thinking for speaking’. In John Gumperz & Steven Levinson (eds.), *Rethinking Linguistic Relativity*, pp. 70–96. Cambridge: Cambridge University Press.
- Stam, Gale (1998) Changes in patterns of thinking about motion with L2 acquisition. In Christian Cavé, Isabelle Guaitella & Serge Santi (eds.), *Oralité et gestualité*, pp. 615–619. Paris: L’Harmattan.
- Talmy, Leonard (1975) Semantics and syntax of motion. In John Kimball (ed.), *Syntax and Semantics* 4, pp. 181–238. New York: Academic Press.
- (1985) Lexicalization patterns: semantic structure in lexical forms. In Shoepen, Timothy (ed.), *Language Typology and Lexical Description* 3, pp. 36–149. Cambridge: Cambridge University Press.
- (2000) *Toward a Cognitive Semantics* 2. Cambridge: MIT Press.
- Whorf, Benjamin (1956) *Language, Thought and Reality*. Cambridge: MIT Press.

Contact information:

Wojciech Lewandowski  
 CLT Centre de Lingüística Teòrica  
 Universitat Autònoma de Barcelona  
 Edifici B, Campus UAB  
 08193 Bellaterra, Spain  
 e-mail: woj(dot)lewandowski(at)gmail(dot)com

## Appendix 1

Examples of the cloze test

1. Hola, ¿cómo \_\_\_ llamas?  
 a) te                      b) se                      c) lo                      d) él
2. – No tengo ni idea. – Yo \_\_\_  
 a) también              b) también no          c) tampoco              d) no sé
3. La madre de mi madre es mi \_\_\_\_  
 a) tía                      b) cuñada              c) abuela              d) sobrina
4. Me parece lógico que no \_\_\_ hablar.  
 a) quiere                  b) ha querido          c) quiera                  d) quise
5. Estaba en paro porque \_\_\_  
 a) lo despidieron                      b) lo han despedido  
 c) lo despidieran                      d) lo habían despedido
6. Ese futbolista es muy joven y \_\_\_ aún un poco \_\_\_  
 a) es/negro              b) es/verde              c) está/negro              d) está/verde
7. Pedro se ha \_\_\_ loco y ha dejado el trabajo.  
 a) hecho                  b) puesto                  c) vuelto                  d) transformado
8. Yo no estoy de acuerdo con \_\_\_ de la casa.  
 a) tal                      b) ello                      c) esto                      d) lo

## Appendix 2

Examples of the judgment task

1. Juan, ven aquí, por favor. (Condition 1, correct sentence)  
 ‘Juan, COME here, please’
2. Gracias por tu visita. ¿Cuándo vas de nuevo a mi casa? (Condition 1, incorrect sentence)  
 ‘Thanks for your visit. When are you going to GO again to my place?’
3. ¿Te apetece venir con nosotros al teatro esta tarde? (Condition 3, correct sentence)

- ‘Would you like to COME with us to the museum this afternoon?’
4. ¿Te apetece ir con nosotros a Ámsterdam el mes que viene?  
(Condition 3, incorrect sentence)  
‘Would you like to GO with us to Amsterdam next month?’
5. A las cuatro vamos desde mi casa a Plaza Cataluña. (Condition 4, correct sentence)  
‘At four we will start GOING from my place to the Catalonia square.’
6. Ponte la chaqueta. Tenemos que venir ya. Juan nos está esperando.  
(Condition 4, incorrect sentence)  
‘Put on your jacket. We have to COME now. Juan is waiting for us.’
7. (Una conversación entre Pedro y Juan en casa de Juan)  
- ¿Qué tal la fiesta que hiciste ayer en tu casa, Pedro? ¿Fue mucha gente?  
(Condition 5, correct sentence)  
(A conversation between Pedro and Juan at Juan’s place)  
- What about the party you made yesterday at your place, Pedro? Did many people GO?
8. (María está en Barcelona y está hablando por Messenger con su amiga Natalia. Natalia está en un pueblo cerca de Barcelona)  
- Hola Natalia, esta noche hago una fiesta en mi casa. ¿Te quieres pasar?  
- ¿En tu casa? Claro que sí, pero vendré sobre las doce.  
(Condition 5, incorrect sentence)  
(María is in Barcelona and she is talking by Messenger with her friend Natalia. Natalia is in a town near Barcelona.)  
- Hello, Natalia, there is a party at my place this night. ¿Would you like to pop in?  
- At your place? Sure, but I will COME about twelve.



**Ulla Vanhatalo, Heli Tissari & Anna Idström**

## **Revisiting the Universality of Natural Semantic Metalanguage: A View through Finnish<sup>1</sup>**

### **Abstract**

The Natural Semantic Metalanguage (NSM) is a method of semantic analysis, used for various tasks mainly in the field of linguistic research. A crucial part of the theory is the set of primes, minimal lexical units that are used to explicate words, cultural scripts and other concepts. Identifying the primes in a new language is an opportunity to reinforce and/or revisit the theory. The remarks presented in this paper resulted from the identification process of the Finnish-based NSM primes. The goal of this paper is to direct attention to some fundamental aspects in the Natural Semantic Metalanguage theory, especially to the relation between the universal language-independent NSM concepts and the English-based NSM. A number of remarks are made on the general system of the primes, as the paper points out issues related to e.g. the number, selection and mutual hierarchy of the primes. The economy and logic of certain prime constructions and the argumentation behind allolexy are discussed as well.

### **1. Introduction**

**The Natural Semantic Metalanguage** (henceforth: **the NSM**) is an approach to linguistic meaning, originated by Anna Wierzbicka in the 1970's, and developed further mainly by Anna Wierzbicka and Cliff Goddard. The NSM is based on three fundamental assumptions: 1) *there exists a natural semantic metalanguage, through which all words in every language can be defined;*<sup>2</sup> 2) *this metalanguage is based on a set of indefinable atom-like words called primes, and a simple grammar;* 3) *these*

---

<sup>1</sup> The manuscript was submitted for the first time in March 2012. After three rounds of refereeing, the published version of this paper owes a lot to the most helpful and patient anonymous referees. We want to sincerely thank Prof. Cliff Goddard for very constructive comments on different versions of the manuscript.

<sup>2</sup> The result of an NSM analysis process is called *explication*.

*primes*,<sup>3</sup> together with their associated grammar, can be identified in every natural language.

There are two main guidelines for identifying primes. Being a prime, a lexical unit must (a) be found in every natural language, and (b) be indefinable via other primes. To put it simply: All words (more precisely, lexical units) in all languages can be divided into two groups: the primes (64, plus some functional synonyms, *allolexes*) and the other words. The words in the latter group can, in principle, be explicated through the primes. As all NSM versions based on different languages are mutually fully translatable, any NSM version can be used for explicating any word in any language.

Although there is nowadays an increasing dialogue between the NSM and other semantic paradigms, there still remain disputed issues. Most importantly, the idea that primes are the core of a universal mental lexicon (along with the idea that most words other than primes are language-specific) has been challenged by many critics (e.g. Murray & Button 1988, see also Wierzbicka's reply 1988; van Driem 2004; Wawrzyniak 2010; Enfield 2002; Geeraerts 2010; McCawley 1983). Not enough attention has been paid to the unavoidable contradiction that the primes have firstly been identified in English, yet they are intended to be (and are considered by their proponents to be) language-independent. Indeed, the relationship between the primes as mental concepts and their manifestation in the world's languages deserves more extensive discussion than has been conducted so far. Moreover, a neutral analytic comparison between the NSM and other methods in the field of semantic analysis is also missing.

In the NSM approach, locating primes in a new language is seen as reinforcing the validity of the general theory of NSM. The primes are now considered to have satisfactory counterparts in more than 30 languages.<sup>4</sup> Most of the findings in these papers seem to support the general theory of primes; however, some criticism has also been presented (e.g. Goddard & Karlsson 2008). An extensive literature already exists on careful

---

<sup>3</sup> The counterparts of the primes in different languages are called *exponents*. If the exponent of certain prime in a certain language based NSM variant has multiple variants, they are called *allolexes*. For examples, see Table 1.

<sup>4</sup> E.g. in Amharic (Amberber 2008), East Cree (Junker 2008), French (Peeters 1994), Japanese (Onishi 1994), Korean (Yoon 2008), Lao (Enfield 2002), Mandarin (Chappell 2002), Mangaaba-Mbula (Bugenhagen 2002), Malay (Goddard 2002), Polish (Wierzbicka 2002), Russian (Gladkova 2010, for the latest set, see the NSM home page), Spanish (Travis 2002), and Thai (Diller 1994).

translations of primes into many languages (e.g. Gladkova 2010; Yoon 2006). However the field test of these prime sets, i.e. their application into actual semantic explications remains less well examined. The vast majority of all NSM explications are made and published in the English-based NSM.<sup>5</sup>

In the Finnish project, the newly invented Finnish primes were tested with respect to their capability to explicate the meaning of certain words (Vanhatalo & Tissari, forthcoming).<sup>6</sup> Many concerns reported in the present paper were exposed only in the translation process of explications (for a note on previous translation problems, see also Peeters 1994: 440). We needed to revisit our initial proposals and make adjustments (for example, in the case of *kind of* vs. *like*), and still some questions remained open.

The aim of the present paper is to revisit some fundamental aspects of the NSM theory based on observations made during the creation of a Finnish version of the NSM. As the remarks are ensuing from the identification work done with the Finnish-based primes, the Finnish primes are presented as a Table in section 2 (in more detail, Vanhatalo & Tissari, forthcoming). We have collected our comments in the following two sections under two closely related main headings: The relation between the universal NSM concepts and the English-based NSM (Section 3), and remarks on the general system of primes (Section 4). Section 5 will discuss the findings.

---

<sup>5</sup> A lot of explications have been done in Polish and Russian (mainly Wierzbicka, also Gladkova), and some in French (Peeters), Arabic and Hebrew (Habib). Only part of the work has been discussed and published in English. The more NSM explications made in languages other than English will be analysed in English, the wider audience will be able to discuss, analyse, criticise and improve the work.

<sup>6</sup> It is to be mentioned here, that the current paper is not the first endeavor to identify the NSM primes in Finnish. In her PhD thesis, Seija Tuovila (2005) published a selected set of NSM primes to be used in explications of Finnish emotion vocabulary. The thesis also included quite an extensive selection of compact NSM explications.

## 2. The Finnish<sup>7</sup> based NSM primes

In the course of the identification process of the Finnish exponents, all the basic combinatorial possibilities of primes (from Goddard 2011c) were translated into the Finnish-based NSM (for examples, see Appendix 1.). After this, some NSM explications were translated into the Finnish-based NSM. These included the verb *promise* (Engl.), the discourse particle *well* (Engl.), the noun *a cup* (Engl.), the noun *God* (Engl.), and the interjections *Psst!* (Engl.), *Pst!* (Pol.), *Shh!* (Engl.) and *Sza!* (Pol.). Some examples are presented in Appendix 2. After the identification process, the NSM Finnish version has been successfully used in numerous explications of cultural scripts, cultural key words, states of emotions, social situations, etc. (Vanhatalo forthcoming).

The set of the Finnish primes is presented as currently<sup>8</sup> seen in “state of the art” NSM research.

**Table 1.** The set of the Finnish primes, version 1.

English (NSM homepage 2013 <sup>9</sup> )	Finnish (Vanhatalo & Tissari, forthcoming)
I	MINÄ
YOU	SINÄ
SOMEONE	JOKU~IHMINEN~HÄN
PEOPLE	IHMISET
SOMETHING~THING	JOKIN~ASIA
BODY	RUUMIS~KEHO
KIND	-LAINEN~-LÄINEN

<sup>7</sup> Finnish, typologically located between fusional and agglutinative language types, is a Uralic language spoken by 5–6 million speakers mainly in Finland. In Finnish vocabulary, many words are created with derivational suffixes, with verbal suffixes in particular being extremely diverse. From the lexical semantic point of view, the Finnish language provides an interesting laboratory, as Finnish is mainly spoken by native speakers in a geographically delimited and linguistically relatively homogeneous area.

<sup>8</sup> With the prime SOMEONE, however, we have made an exception by using the earlier NSM version (Wierzbicka 1972). For more discussion, see 4.3.

<sup>9</sup> The current table of NSM primes has been under development for decades. In the very first set (Wierzbicka 1972), as few as 13 primes were presented. Later on, there have been continuous updates (both withdrawals and additions) to the set of primes, the total number being currently (2014) 65. The very last change in the set of primes has been adding DON'T WANT as an individual prime (August 14th 2014 update on the NSM homepage). As the case of DON'T WANT would require deeper analysis than is allowed by the publishing process of this paper, we use the version 2013 with its 64 primes as our point of reference. Goddard describes the history of primes in terms of three generations (2011a).

PART	OSA
THIS	TÄMÄ~SE
THE SAME	SAMA
OTHER~ELSE	TOINEN~MUU
ONE	YKSI~ERÄS
TWO	KAKSI
SOME	JOKIN~JOKU~MUUTAMA
ALL	KAIKKI
MUCH~MANY	PALJON~MONI
LITTLE~FEW	VÄHÄN~HARVA
GOOD	HYVÄ
BAD	PAHA
BIG	ISO~SUURI
SMALL	PIENI
THINK	AJATELLA
KNOW	TIETÄÄ
WANT	TAHTOA
FEEL	TUNTEA~TUNTUA
SEE	NÄHDÄ
HEAR	KUULLA
SAY	SANOA
WORDS	SANAT
TRUE	TOSI~TOTTA
DO	TEHDÄ
HAPPEN	TAPAHTUA
MOVE	LIKKUA
TOUCH	KOSKEA
BE (SOMEWHERE)	OLLA (JOSSAIN)
THERE IS	OLLA (OLEMASSA)
HAVE <sup>10</sup>	OLLA (OMISTAA)
BE (SOMEONE/SOMETHING)	OLLA (JOKU/JOKIN)
LIVE	ELÄÄ
DIE	KUOLLA
WHEN~TIME	MILLOIN~JOLLOIN~SILLOIN~ AIKA
NOW	NYT
BEFORE	ENNEN~AIKAA SITTEN~AIKAISEMMIN
AFTER	JÄLKEEN~AJAN KULUTTUA
A LONG TIME	KAUAN (AIKAA)~PITKÄN AJAN
A SHORT TIME	VÄHÄN AIKAA
FOR SOME TIME	JONKIN AIKAA
MOMENT	HETKI
WHERE~PLACE	MISSÄ~JOSSA~PAIKKA

---

<sup>10</sup> In the last versions, the prime HAVE has been replaced with the prime BE (SOMEONE'S)

HERE	TÄSSÄ~TÄÄLLÄ
ABOVE	PÄÄLLÄ
BELOW	ALLA
FAR	KAUKANA
NEAR	LÄHELLÄ
SIDE	PUOLI
INSIDE	SISÄ-
NOT	EI
MAYBE	EHKÄ
CAN	VOIDA
BECAUSE	KOSKA~VUOKSI~TAKIA
IF	JOS
VERY	HYVIN~ERITTÄIN
MORE	ENEMMÄN~LISÄÄ (and ENÄÄ)
LIKE~AS~WAY	NÄIN~KUTEN

### 3. The relation between the universal NSM concepts and the English based NSM

One of the very first observations in the course of the Finnish project was that the exponents of the primes involve many kinds of polysemy. Polysemy may occur in two contexts at least: 1) The English-based NSM exponents may be polysemous, e.g. BAD or LIVE; 2) Any other language based NSM exponents may be polysemous, e.g. the Finnish TEHDÄ which is only allowed to function as ‘to do’ while ‘to make’ is prohibited although the verb has both meanings. From the point of view of a new NSM version, the crucial question is whether that polysemy is something that just happens to occur with the English version, or should the same explicative functions be found in other language based NSM variants, using some other lexical patterns. These questions led us to consider the role of the English language in the theory of NSM, to be more precise, the relation between the universal NSM concepts and the English-based NSM, the first one referring to the universal language-independent concept of primes, the second one referring to just another language variant of NSM.

It is not fully unambiguous to define the very core of NSM in a theory that claims itself language independent yet uses English-based primes as a metalanguage. This question warrants further attention. We will enlighten this concern in the following by analysing and discussing our findings.

The basic assumption in the NSM is that the primes are strictly identified by the proposed basic combinatorial possibilities (from Goddard 2011c, see Appendix 1). These sentences set the grammatical restrictions

for every prime, and the NSM explications can be made with these combinations only. The basic combinatorial possibilities are to determine that every prime is only used within the foreseen lexical relations. In other (and more conventional) words, only one sense of a certain word from a given language is selected to serve as an NSM prime. Being selected and confirmed as a prime, the lexical item is supposed to be an independent and indefinable (in the sense that it cannot be rephrased) meta-level unit. It is supposed to behave in the NSM explications cleanly without interference from the other lexical roles and relations it has in its normal use in the natural language from which it comes.

The situation is still not clear. Many problems caused by polysemy of the exponents of the NSM primes in particular languages are already well recognized. Within the English-based NSM, exponents of primes like FEEL and KNOW are indisputably polysemous (e.g. Goddard & Wierzbicka 1994: 31–32). The solution has been to describe the use of these words in detail in order to ensure both that the appropriate sense is used and that the appropriate lexical item is searched for in other languages. In some cases, cross-linguistic comparison has prompted a re-evaluation of certain primes. For example, the NSM understanding of THINK was challenged by the evidence from Swedish (Goddard & Karlsson 2008). The exercise showed that the semantic prime THINK has a more restricted grammar than the word *think* in ordinary English, but this had not been previously noted in the basic combinatorial possibilities. The situation with the English verb *think* and the prime THINK may be similar to the one with the English adjective *bad* and the prime BAD (see later in this section).

Based on the historical fact that the NSM primes were first used extensively in English, it is always possible that our understanding of them carries some unrecognized hidden features that are specific only to the English language. The only way to find out those hidden features is to carefully look at every set of primes identified in new languages. Even though already having the strict grammatical and contextual restrictions, the original exponents of the NSM primes are not yet free from ordinary language-dependent lexical relations, mainly polysemy. These relations, however, may only become visible through continuing the process of transposing explications from one NSM variant to another. The following examples from the Finnish language enlighten this question.

*The case of ONE.* The prime ONE (like TWO) is identified as purely a quantifier (Wierzbicka 1996: 44). In published explications, however, this

is not always the case, as can be seen in a part of the definition of *a cup* (Goddard 2011a: 229–230): *someone can hold<sub>[m]</sub> one in one hand<sub>[m]</sub>* (focus on the first use of *one*). The Finnish-based NSM does not support this kind of use of the prime ONE, as can be seen in the translation *ihminen voi pitää<sub>[m]</sub> sellaista yhdellä kädellä<sub>[m]</sub>*. In the same vein, expressions like *the other one* are being translated as *tuo toinen*, literally, ‘that other’. Or, in some cases, the translation could be even *tuo toinen tällainen*, literally, ‘that other this kind’, using the prime KIND. According to Cliff Goddard (personal discussion November 2011), this kind of secondary use of the quantifier ONE could in principle be replaced with expressions like *one something of this kind*, even though it could make the explication harder to understand.

*The case of SOMETHING~THING.* The Finnish language seems to be lacking the overall concept for the prime THING, which may well refer to concrete, as well as abstract, objects. The Finnish word *asia* is normally used about abstract subjects, while *esine* refers to concrete nouns. The latter has certain restrictions, as it is only used for relatively small and non-living objects. A house or a cat could not be referred as *esine*, while they usually do not belong to the abstract category of *asia* either. The solution in most of the cases is to use just the prime JOKIN (‘something’), to refer to the object in question.

*The case of LIVE.* Even though the English verb *to live* is used in contexts like *they are living together* or *they live in the midst of a forest*, the prime LIVE does not include this type of adjunction, as pointed out by Wierzbicka (1996: 86–87). Making the difference in the English-based NSM seems still to be quite challenging, as seen even with the basic combinatorial possibilities: *many people live in this place, someone lives with someone else*. It is not always clear whether the sentences deal with Finnish *elää* ‘to live, to be alive’ or *asua* ‘to live, to stay, to reside’.

The point with the examples above is that the apparent polysemy of the primes under consideration may be specific to the English-based NSM exponents only, not necessarily to the universal NSM primes.

*The case of MORE.* From the Finnish point of view, the NSM prime MORE has, in addition to its main use (e.g. *joku haluaa enemmän / lisää* ‘someone wants more’) another specific use that we see as a separate meaning. In the list of the NSM primes, only MORE is listed, however, in practice *anymore* is used as well. Regardless of the fact that the word *anymore* in English contains the element *more*, the following two basic combinatorial possibilities *ei elä enää* ‘not living anymore’ and *ei enää*

*kuten tämä* ‘not like this anymore’ seem to deal with a different semantic concept. Namely, the English *more* refers to quantity, while *anymore* deals with (the length of) time. Quantity and time are different measures, and they cannot be combined at the level of primes without logical problems (see also the time expressions in Section 4). According to Cliff Goddard (personal discussion November 2011), one reason for linking *more* and *anymore* is the dynamic “heading forwards” feeling both of them have. The authors of this paper doubt whether this kind of consideration is sufficient, and moreover, whether the dynamicity is a property of the NSM prime as such or just a property that goes with the English language. As another “time-quantity” question arises in connection with the prime FOR SOME TIME, we suggest that both these primes deserve to be re-thought by the originators of the NSM. There is a certain mismatch in the identification of the primes.

*The case of BAD.* BAD is one of the very fundamental primes in the NSM, and it was among the most challenging translation tasks in the Finnish-based NSM project, hence it will be discussed in more detail.

The Finnish language has two main candidates for the exponents of BAD: PAHA and HUONO. The Finnish *paha* in many contexts comes close to English ‘evil or immoral’, while *huono* usually describes something as ‘low in quality’.<sup>11</sup> Both of them are opposites of *hyvä* ‘good’. In addition to *paha* and *huono*, there is also an adjective *tuhma* ‘incorrectly behaving, naughty’, the opposite of *kiltti* ‘correctly behaving’, both meanings expressed in English with *bad* and *good* (*good girl, bad boy*). As there are no obvious differentiated English counterparts for the Finnish *paha*, *huono* and *tuhma* (regardless of which of them is to be chosen as the official Finnish exponent of BAD), we have a good reason to believe that the semantic content of all these words are expressed by the NSM prime BAD.<sup>12</sup> As *tuhma* has remarkably different range of uses, we leave it aside from most of the further considerations (for a further discussion of ‘good girls, bad boys’ see Wierzbicka 2004).

The basic combinatorial possibilities can all be translated into the Finnish-NSM by using one or other of the candidates PAHA and HUONO,

---

<sup>11</sup> Explaining the Finnish *paha* with the English words *evil* or *immoral* does not actually help us to understand the very nature of this word, as both *evil* and *immoral* have their own complex meanings. Counterparts like these are to be taken with caution.

<sup>12</sup> An interesting aspect to BAD comes from e.g. the *Merriam-Webster Dictionary*: the first definition of English *bad* is quite close to the Finnish *huono*: “failing to reach an acceptable standard”.

and sometimes with both but with separate meanings (see Appendix 1; for details Vanhatalo & Tissari, forthcoming). One can also contrast between *paha* ‘bad, evil’ and *huono* ‘bad, low-quality’ in Finnish, saying, for example, about an alcoholic: *hän on huono ihminen, mutta ei hän paha ole* ‘s/he is a bad (not good) person, but s/he is not bad (evil)’. In the same vein, one can contrast ‘evil’ and ‘non-properly behaving’, saying, for example, of an poorly behaved student: *hän on tuhma, mutta ei paha* ‘s/he is bad (non-properly behaving), but not bad (evil)’. Furthermore, *ruoka ei välttämättä ole huonoa, vaikka se maistuisi pahalta* ‘food is not necessarily bad (in quality) even if it tastes bad’.

These examples led us to think about the true universality of the prime BAD. Is the very nature of BAD rather ‘paha’ or ‘huono’ – or something else, from some other language perspective? From the philosophical point of view the question is what the fundamental ideas of GOOD and BAD are (Goddard & Wierzbicka 1994: 47). Wierzbicka (1994: 496–497, 1996: 51–54) discloses her (and other researchers’ as well) earlier thoughts on these evaluative adjectives, mainly connecting the primes GOOD and BAD (which were not presented in the first set of primes Wierzbicka 1972) with the older prime WANT. She describes the problems with linking GOOD and WANT (the earlier explication of the English word *good* was ‘what someone wants’, and *bad* respectively ‘what someone doesn’t want’), and comes to the conclusion that WANT should not be linked with either of the evaluators. The main reason for breaking the older connection with the evaluators is the simple fact that sometimes one who wants can want something bad. Wishes and desires are subjective, while primes should be objective. Furthermore, even though people may not share their understandings of what is GOOD and what is BAD, they do agree that those concepts do exist (Wierzbicka 1996: 52).

Finnish is not the only language in which a problem arises with BAD. Russian, for example, has even more varieties to choose from: *ploxoj*, *durnoj*, *zloj* and *nexoros(hattu)ij* (Goddard & Wierzbicka 1994: 47, Gladkova 2007: 58). Studies on Mandarin Chinese (Chappell 1994: 142, referred by Wierzbicka 1996: 53), show how the primes GOOD and BAD are semantically asymmetrical, the exponent of BAD being narrower than the exponent of GOOD. According to Chappell (1994: 142), the Mandarin *huài* (exponent of BAD) seems to have some similarity with the Finnish *paha*, namely the aspects of immoral, nasty or evil. Wierzbicka explains the Mandarin variation as cultural rather than semantic (Wierzbicka 1994:

497), but for us the case is less clear. The use of Malay *buruk* (exponent of BAD) is also limited (Goddard 2002: 132).

Earlier in this section we referred to the consequences that research on the Swedish TÄNKÅ caused to the identification of the prime THINK. Yet there is a significant difference between the primes THINK and BAD: While THINK is a mental verb with highly specific syntactic features (complementation possibilities), BAD is used as an attribute or a predicate with less versatile syntax and much “meaningful” semantics. Nonetheless, we would encourage the NSM theory to sharpen the identification of BAD by distinguishing somehow between uses where English *bad* amounts to ‘evil, immoral, nasty, unpleasant’, on the one hand, and uses in which it conveys something like ‘low in quality’, on the other. Our suggestion is that the current prime BAD applies only to the ‘evil, immoral, nasty, unpleasant’ uses, while uses of *bad* that convey ‘low in quality’ are not semantically primitive but can be explicated – perhaps by using the negated version of the exponent for GOOD. The latter expression, furthermore, would include all other possible ‘not-good’ uses.

Taken together, the examples of polysemy<sup>13</sup> presented above disclose the very fundamental question about the true universality of the current definition of the NSM primes. It is not always clear what we are actually dealing with: polysemy of the universal language-independent NSM primes or polysemy of the English exponents of the NSM primes. The question is crucial. According to Goddard (personal discussion November 2011), there is no overriding need to avoid inter-NSM (or inter any-language-based-NSM-version) polysemy. The burning question to us has been to make explicit the difference between ultimate inter-NSM polysemy and polysemy in the English-based NSM version, the first one having obviously more fundamental consequences for NSM based languages other than English than the latter one.

#### 4. Remarks on the general system of primes

This section presents some remarks on the general system of primes. As the set of the NSM primes is very compact and internally closely related, also

---

<sup>13</sup> Notably, polysemy is involved on the Finnish side, too (Vanhatalo & Tissari, forthcoming). As seen with English examples, sometimes the question about possible polysemy remains hidden until it pops up in a problematic translation process between NSM variants based on two different languages.

the questions raised up in this section and in the entire paper are inherently tied together.

#### 4.1 The number, selection and mutual hierarchy of the primes

The question about the optimal number of NSM primes has been discussed since the early days of NSM. The number of primes has gradually increased over the past 40 years from 13 to 64, and some initially introduced primes have been removed (e.g. IMAGINE and WORLD, Wierzbicka 1972). Even though the primes are seen as the solid core of NSM, the general tendency seems to favor new primes slowly coming into the set.

*The case of BE (SOMEWHERE), THERE IS, HAVE, and BE (SOMEONE/SOMETHING).* The Finnish-based NSM challenges the existence of some of the current primes. In the Finnish set of primes, the verb OLLA ‘to be’ serves for four NSM primes, as it works for locational, existential, possessional or specificational verbal primes. The primes can still be distinguished by grammatical properties, as noted in the basic combinatorial possibilities. Some of these primes might be explicated with the other primes.

*The case of I and YOU.*<sup>14</sup> Related to the expansion of the primes, a question arises about the internal relations between the primes. One can wonder whether all the primes are mutually equal or is it possible that some are e.g. more “primary” in some sense than the others. There seems to be no way to understand the prime YOU without first understanding the prime I, SINÄ without first understanding MINÄ. Arguably, YOU is always understood in relation to I, which is definitely something indefinable and could be considered to be a quintessentially primary prime.

*The case of THIS and HERE.* Another example of a possible hierarchy is provided by the Finnish TÄSSÄ~TÄÄLLÄ ‘here’. These forms are literally inner and outer locative cases of the pronoun *tämä* ‘this’, which raises a question about the relation between the primes THIS and HERE. As we cannot have TÄSSÄ~TÄÄLLÄ ‘here’ without TÄMÄ ‘this’, one of them seems to be more primary than another. The fact that in English HERE and THIS are different lexemes should not guarantee that they are different NSM primes. It should be carefully studied whether one of them

---

<sup>14</sup> This remark is not specific to Finnish.

could be explicated by other primes. In that case, *HERE* might be just an allolex of *THIS*.

An indirect argument in support of the idea of recognising more and less primary primes comes from the history of the development of the prime set itself, as many of the early explications were made without certain currently proposed primes. NSM work over past decades has shown an unbelievable ability to produce sound explications with a number of different prime sets. Some of the current troubles with, for example, the apparent polysemy of some exponents of primes might result from an over-development of a compact and highly interrelated system. In other words, the number of primes in the current set is possibly too high. As all natural languages have their own issues with various lexical relations, we come again to the point that it is necessary to keep the various variants of NSM tightly tied to the universal NSM primes, and to avoid any unnecessary contact with, and contamination from, the English version (and lexical relations in the English language).

## 4.2 Economy and logic of primes

*The case of time expressions.* The primes expressing time raise questions about the logical consideration behind certain primes. The main question is whether the primes expressing time should be individual combinations as they are now or should they just be combined by separate primes as needed in use. Many languages have specific lexicalizations for expressing ‘for a long time’ and ‘for a short time’, with no morphological element corresponding to ‘time’ being included at all. This has been used in some NSM work as an argument in favour of the unitary nature of *A LONG TIME* and *A SHORT TIME*, i.e. against the idea that these meanings could be composites of *TIME* and a quantifier (such as *MUCH* and *LITTLE*). Actually the same feature is found in Finnish, where we have the alternative expression *kauan* ‘for a long time’. It is still not really valid to take such considerations into account when deliberating about the ultimate level of the NSM primes if these kinds of features are specific only to some languages.

Even though *SOME + TIME* wouldn’t work for the expression *for some time* (as the durational aspect would be missing without the word *for*), *LITTLE + TIME* might work for *a short time* or *MUCH + TIME* for *long time* in combinations like *a long time before*, *a short time after* (Goddard 2011c chart).

*The case of SOME.* According to the NSM policy, *some* in its semantically primary sense cannot be used with uncountable nouns (i.e. expressions like *some water* and *some time* are not allowed). The reason for this kind of restriction is that SOME in that sense is definable through other primes, *not much not little*. According to Cliff Goddard (personal discussion November 2011), the case of SOME is still currently not quite clear, and the identification of the prime may be changed in the future. But yet *some* still appears combined with the most uncountable expression in the independent prime FOR SOME TIME. One might then wonder whether it would be more useful to make this *some* as the actual prime SOME (allowing it to be used with any uncountable noun) and forget the narrow prime FOR SOME TIME.

*The case of KIND and LIKE.* The primes KIND and LIKE in Finnish draw attention to the surprisingly close relationship between these two primes and the very nature of each of them. The connection can be seen with English-based NSM expressions like *I have not seen someone like this* vs. *I have not seen someone of this kind*, which both could be translated to the Finnish-based NSM as *en ole nähnyt ketään tällaista*. Another example would be *someone says something like this* vs. *someone says something of this kind*, both could be translated into the Finnish NSM as *joku sanoo jotain tällaista / joku sanoo jotain näin*. Based on the similarities between these expressions, we could provocatively ask whether KIND and LIKE are truly different NSM primes or are they just English allolexes of one single mental prime. Even though cases like these are in minority in all of the uses of the prime KIND, the question of possible inter-NSM polysemy or synonymy itself is theoretically interesting, and worth deeper general discussion. Wierzbicka refers to the relationship between KIND and LIKE (1994: 494), but based on our evidence, we cannot fully agree with her arguments on the mutual independence of these concepts.

*The case of KIND.* Having a look at KIND independently, the English noun *kind* (and the English-based NSM prime KIND) is a much “stronger” and more wide-ranging concept compared with the Finnish suffix *-lainen~-läinen* (and thus also the prime -LAINEN~-LÄINEN). For example, it is quite hard to translate categorizing expressions like *natural kinds* into Finnish without a specific term *laji* (‘species’). One might even consider taking *laji* (‘species’) as an allolex of the prime -LAINEN~-LÄINEN.<sup>15</sup>

---

<sup>15</sup> The Finnish *laji* ‘species’ is actually the etymological origin of the suffix *-lainen ~ -läinen*, as was kindly pointed out by one of the anonymous referees.

### 4.3 Grammatical number of the primes

The question about the grammatical number of primes is mostly involved with the much discussed prime PEOPLE, much less attention has been paid to the prime WORDS. These two primes are the only ones having no singular form. From the Finnish point of view, the grammatical number of these primes seems to be somewhat problematic. It is hard to see why we should use the plural forms IHMISET ‘people’ and SANAT ‘words’ instead of the simple singular forms IHMINEN ‘human being, person’ and SANA ‘word’. Logically, the singular form would be the primary one.

*The case of PEOPLE.* The quite extensive discussion about PEOPLE is closely related to SOMEONE and the status of the word *person* in the NSM context (e.g. Goddard & Wierzbicka 2002: 44–45, 79; Goddard 2002: 20; Goddard & Wierzbicka 1994: 33). Goddard and Wierzbicka argue against having PEOPLE decomposed as the plural of SOMEONE or PERSON by saying that PEOPLE is restricted to humans while SOMEONE is not. Nonetheless, in the majority of the cases, isn’t the identification of SOMEONE exactly what we understand with human beings, even though the English language does not have an optimal word for it? Still we have to admit that in some cases, there can exist “someones” who cannot get the human label, namely, gods, aliens and other kind of beings (e.g. Habib 2011). The Finnish version of NSM ended up with adding IHMINEN ‘human being, person’ and HÄN ‘s/he’ as allolexes of SOMEONE.

The main reason for adding IHMINEN ‘human being, person’ and HÄN ‘s/he’ as allolexes of JOKU ‘someone’ is to ensure that Finnish NSM explications will stay clear and understandable. We are well aware of the slight but acceptable oddness of the formulation *this someone* in the English-based NSM, and regret the similar formulation *tämä joku* ‘this someone’ (with grammatical allolexes *tämän jonkun* (genitive), *tätä jotakuta* (partitive), *tänä jonakuna* (essive), etc.) would sound too odd in the Finnish-based NSM. Including the allolex IHMINEN ‘human being, person’ into the Finnish-based NSM does not go against the original thought of the NSM. As we still cannot see good enough universal reason to make the general distinction between singular and plural forms (even though this may be the case within the English language), we would suggest to consider removing the prime PEOPLE entirely from the NSM set of original primes, and adding allolexes PERSON and PEOPLE to the prime SOMEONE in the English version of NSM.

*The case of WORDS.* Another suggestion is related to the grammatical number of WORDS. According to Cliff Goddard (personal discussion November 2011), there is no fundamental reason to have the prime WORDS in the plural form. The prime is presented in the plural because it is most frequently used in plural (see Goddard 2011b). Thinking about the general philosophy of the NSM, however, we suggest considering having the prime in singular, as it is simpler than the plural. This agrees with Sinnemäki's (2011: 16) definition of complexity, one characteristic of which is that there are a number of elements in a structure.<sup>16</sup>

#### 4.4 Argumentation behind the lexical allolexes

The lexical relations involved with the identification of primes are not restricted to polysemy. The cases of synonymy led us to think about both diachronic and synchronic lexical variation. The lexical meaning does change over time, space and context and it may have some effect on the identification of NSM primes.

The fact that the current written Finnish includes features from two main dialects, had some implications for the Finnish NSM project as well. Some of the synonyms that once served as counterparts in two different dialects, may have developed differentiated meanings after the formation of a common written Finnish language. In the study on Finnish primes (Vanhatalo & Tissari, forthcoming), it was not always clear what the best Finnish candidate for prime status is, and whether or not the erstwhile synonyms should both be taken into account as NSM allolexes (e.g. Western *iso* vs. Eastern *suuri* 'big', *haluta* vs. *tahtoa* 'want').

The term *allolex* is used in the NSM context to mean a variant of a prime. The allolexes have the same identification ('meaning') as the prime does, and they are used for grammatical or collocational reasons. Some of the allolexes are just morphologically different forms of the primes (*minä* 'I', *minun* 'my'), some can be different lexemes (*toinen*, *muu* 'other, else'). It is not always straightforward to decide whether some words should or

---

<sup>16</sup> Both Goddard and an anonymous reviewer seem to be right, though, about the frequency of the plural *words* as compared to the singular form *word*. The 450 million-word *Corpus of Contemporary American English* (1990–2012) attests 78656 instances of the type *word* as against 98366 instances of the type *words*, while the 100 million-word *British National Corpus* (late 20th century) attests 18707 instances of the type *word* as against 23632 instances of the type *words*.

shouldn't be regarded as allolexes of certain prime, as seen in the following examples.

*The case of BIG.* The Finnish variant of NSM has allolexes ISO~SUURI, because they are claimed not to have remarkable meaning differences (this has been noticed in dialectological studies as early as Nirvi (1936: 30–32), although the situation may have changed since then). Further studies on these lexemes would be most welcome, as at least some specialized usage can be found, e.g. *hän on iso mies* 'he is a big man' vs. *hän on suuri mies* 'he is a great man', the first one referring to physical size while the latter one rather describes a mental or social feature. Interestingly, Wierzbicka does not seem to pay any attention to the English variation *big ~ large*, neither does she argue why it is just the English *big* that has been chosen to be the (only) exponent of BIG (1996: 54–55). According to Cliff Goddard (personal discussion November 2011), *large* is narrower in its range of use and can be defined through BIG.

*The case of WANT.* The search for the Finnish exponent for WANT gave us another pair of synonyms – HALUTA and TAHTOA, the meaning difference of these verbs possibly correlating with the difference between the nouns *desire* and *will* (respectively).<sup>17</sup> The differences between these two candidates led us to an interesting question about the deepest aspects of wanting. According to Finnish dictionaries, *haluta* roughly means 'someone's action while trying to fill some need or to make some wish true', while *tahtoa* is 'determined and purposeful action when trying to reach some goal'. The verb *haluta* may have a more primitive and even sexual flavour (*haluan sinua/sinut* 'I want/desire you', the Finnish object can be either partitive or accusative), while *tahtoa* may be more cultivated and conscious and even controlling or controlled action.<sup>18</sup> The verb *haluta* may have the component of owning and using, which *tahtoa* does not have at least in that degree. The nouns connected to the verbs are *tahto* and *halu*, the latter one is intuitively quite close to *himo* 'lust', which is the base for the verb *himoita* 'to desire, to lust (after)'. The differences come nicely visible through compound words like *tahdonvoima* 'willpower', *ruokahalu* 'appetite', or sayings like *luja tahto vie miehen vaikka läpi harmaan kiven* 'a strong will takes a man even through a grey stone'. The verbs *haluta* and *tahtoa* can be used as alternatives in many contexts, e.g. *tahdon/haluan*

<sup>17</sup> We thank an anonymous referee for pointing out this correlation to us, although it may suffer somewhat from circularity of reference: *The Oxford English Dictionary* defines the noun *will* partly in terms of the noun *desire*.

<sup>18</sup> Sound research results on the semantics of these synonyms are lacking.

*muuttaa kaupunkiin* ‘I want to move to a city’, *tahdon/haluan ruokaa* ‘I want to get food’. The verbs still do have many different uses; e.g. the official question and answer used when a couple is getting married: *Tahdotko – osoittaa – rakkautta –?* ‘Do you want to – express – love –?’ and *Tahdon* ‘I do’, to express the very conscious step being taken.

Without a deeper research (which would definitely be very much welcome), we propose that *tahtoa* might express more neutral wanting than *haluta*, and *haluta* could possibly be defined through *tahtoa*. Our current suggestion for the exponent of the prime WANT is thus TAHTOA.

The cases above should lead us to think about the lexical changes taking places around the border between a language and a dialect. The fact that this border is undoubtedly vague and in constant change, supports the idea of at least a certain level of vagueness of primes as well. Although the various cases with their roots in dialectology are not necessarily reported in this paper, this aspect might provide hints for a general discussion about the nature of the NSM primes – and the border between a language and a dialect. Furthermore, still related to the history of lexical changes, there are some cases where the current spoken language differs remarkably from the written one, e.g. the case of the Finnish pronoun *se* ‘it’ widely referring to humans.

*The case of ONE.* The prime ONE can also be viewed through its allolexes. In the Finnish set of primes, there are two possible candidates for ONE, namely YKSI and ERÄS, the first one being a numeral and marking the number 1, while the latter one is more like the indefinite article *a/an* in English. According to the traditional Finnish grammar rules, *eräs* means a referent known to the speaker but not to the hearer while *yksi* refers to something more vague. The definitions of *eräs* and *yksi* have recently been softened to correspond more adequately to linguistic reality. Nowadays both are accepted as equivalents when used as determiners.

*The case of KNOW.* Certain kind of synonymy can also be recognised with the use of English KNOW with the basic combinatorial possibility *joku tietää jonkun toisen ihmisen (hyvin)* ‘someone knows someone else (well)’. When in Finnish *joku tietää* or *tuntee hänet*, the latter verb (which can be translated ‘feel’) is implicating deeper knowing. In Finnish, the verb *tietää* conveys knowing someone just a little (e.g. by name or face), while adding the particle *hyvin* ‘well’ means a little deeper degree of knowing. To express knowing someone truly well, one must use the verb *tuntea* (which can be translated ‘feel’), and one can emphasize the meaning with the particle *hyvin* ‘well’. For example, *tiedän naapurini kerrostalossa* ‘I know

my neighbours in the block of flats (by face)', *tiedän hyvin erään ihmisen joka pelkää lentämistä* 'I know someone well who is afraid of flying', *tunnen työtoverini* 'I know my work mate well', *tunnen hyvin itseni* 'I know myself very well'. The polysemy of the English exponent of the prime KNOW has been deeply discussed in Wierzbicka 1992, but studies around KNOW have been started again (in particular, "knowing someone" will possibly be removed from the set of potential combinations; Goddard personal discussion in March 2012).

#### 4.5 Grammatical allolexy of primes

The complex Finnish nominal case system causes a very large number of grammatical allolexes<sup>19</sup> (Goddard & Wierzbicka 2002: 20) for most of the substantive, verbal, adjective and numeral primes. In the context of NSM, not very much attention has been paid to the significance of extensive morphological phenomena of this kind.

The reason for this is that only a part of the English prepositions are allowed in the English based NSM – as some of them can be explicated with primes. Consequently, not all of the nominal cases (14–15 altogether) should be allowed in Finnish-based NSM explications, however, we find it notoriously challenging to define which ones. For example, can the Finnish elative and illative forms be used while *from* and *to* are not used in the English explications (instead, the formula "something was in one place/now it is in another place", is used)? The key question is: are these kinds of restrictions really driven by the universal language-independent nature of the NSM? It could also be logical to argue against the use of locative cases by noting that they are paraphrasable, i.e. not primitive (Cliff Goddard, personal communication in September 2014).

*The case of the locative cases.* The locative cases combined with spatial expression provide an enlightening example of this consideration. Let us have a look at BELOW (ALLA – *alle* 'to' + 'below', *alla* 'at' + 'below', and *alta* 'from' + 'below'), FAR (KAUKANA *kauas* 'to' + 'far', *kaukana* 'at' + 'far', and *kaukaa* 'from' + 'afar'), and NEAR (LÄHELLÄ *lähelle* 'to' + 'near', *lähellä* 'at' + 'near', and *läheltä* 'from' + 'near'). Use

---

<sup>19</sup> We use the term *grammatical allolexy* to describe the different case forms of the primes (e.g. MINÄ (I) > *minä* (nominative), *minun* (genitive), *minua* (partitive), *minuna* (essive)), to mark the difference from the general NSM allolexy referring different lexemes (e.g. MUU~TOINEN (OTHER)).

of the cases add to the primes a temporal before-after perspective. So *kauas* ‘to somewhere far’, for example, implies that ‘after some time’ the moving person/thing will be ‘far from the place where it was before’. The English NSM would not allow directional ‘to’ and ‘from’ (noticed by Cliff Goddard, personal discussion in March 2012). We would suggest the NSM approach to consider the pros and cons of such practice.

*The case of the abessive case.* Another example comes from verbs. Problems occur with the prime CAN (VOIDA) in connection with abessive form (-*matta*/*-mättä*). The relatively widely used formulation ‘someone can’t not do something’ sounds quite awkward in the English-based NSM, and the strangeness of the Finnish *joku ei voi ei tehdä jotain* ‘someone can’t not do something’ is definitely at least as notable. In this case, we have decided to follow the NSM rule that exponents of the primes in different languages may look grammatically complex, even if expressing ‘simple’ meanings, and suggest the use of third infinitive abessive form: *joku ei voi olla tekemättä jotain* ‘someone can’t be without doing something’. One may ask why the English-based NSM does not use the expression *someone has to do something*, instead of *someone can’t not do something*. The English verb *have to* was in fact used in earlier NSM explications, but as pointed out by Goddard (2014), these two expressions (i.e. *can’t not do* and *have to do*) have different meanings. This can be seen, for example, in the context of someone who cannot prevent herself grinding her teeth at night: we could say that this person ‘can’t not grind her teeth at night’, but not that ‘she has to grind her teeth at night’. Carrying this argument over to the Finnish-based NSM, we can agree to recognise that there is a meaning difference between expressions like *ei voi olla purematta hampaitaan* ‘can’t not grind her teeth’ and *täytyy purra hampaitaan* ‘has to grind her teeth’.

The discussion about grammatical allomorphs of primes and the semantics of grammar raises a number of open questions about the general system of primes. For example, do the constructions like *I think* and *I’m thinking* have any difference that should be carried over to Finnish? The NSM grammar allows both expressions while generally trying to minimise the use of progressive. It seems that the current version of NSM may favor some languages by allowing or prohibiting certain type of complexity. With NSM variants other than English, it is not always clear what we are actually dealing with: legal allomorphs or illegal extended meanings or lexical elements other than primes. Without further studies on grammar, we

do not know whether e.g. a certain Finnish case has too much or too complex “meaning” to be freely used in NSM explications.

## 5. Discussion

The goal of this study was to revisit some basic principles of the NSM approach by using the exploration of the Finnish based NSM primes as the test bed. Sections 3 and 4 presented considerations on the relation between the universal NSM concepts and the English based NSM and, closely related, some remarks on the general system of primes.

The most important finding in our work with the Finnish prime system was that whenever the NSM primes are deployed in even the simplest explications in the English-based NSM (or in any other NSM version), various lexical relations specific to English (or to other languages) are immediately involved. Even the basic combinatorial possibilities referred to in this study (drawn from the Goddard 2011c chart) may be inexact or exclude some crucial elements. Most of the hidden lexical relations become visible only through a translation process to some other-language-based NSM (as with the case of THINK in Goddard & Karlsson 2008).

Our finding is important because attempts to find a convenient and flexible set of allolexes in any language based NSM version can always be hampered by claims like “this is not universal, this cannot be found in all languages”. It should be remembered in this context that the allolexes, lexical relations or grammatical aspects specific to the English-based NSM variant need not be found in other languages (even though some patterns may do so). Every language based NSM version has many allolexes or lexical relations specific to that particular language, which may not be found in any other language. The only requirement for these language-specific allolexes (with their possible lexical relations) is that they cannot be contradictory to the universal NSM primes – an emergent concept that is never 100% finished itself.

Partly due to the variation in allolexes, it is obvious that the NSM versions in different languages may differ in their flexibility and capability in analysing processes. One of our goals in the work with the Finnish NSM has been to make the Finnish-based NSM explications sound real, understandable and acceptable. We feel that at least some parts of the Finnish-based NSM explications were more natural than the originals, and the reason for this may be ignoring some English-specific lexical relations.

This remark should be very welcome, as it proves the vitality of NSM as a universal concept.

The theoretical risk of ending up with an NSM which may not be fully coherent but which has a slightly different toolbox for every language should not obstruct practising and developing the method. There are more pros of having more accurate semantic analysis than cons of possibly compromising with some aspects of the original ideas of NSM. While a critical approach to every method is important, one should be able to compare the shortcomings of the NSM with the ones of other analysis methods. Semantic analysis is among the most challenging tasks in the field of linguistic research, and this far there have not been too many perfect methods available.

Interestingly, even though the official policy regarding the allolexes in the English based variant of NSM is quite strict, the practice has more variation. This can be seen in the original English based explications where some non-primes are used without the molecule<sup>20</sup> mark and with no commentary or other indication that they are intended as allolexes. E.g. the explication of *a cup* (Goddard 2011a: 229–230): *stuff* in the phrase *these other things are made of<sub>[m]</sub> the same hard<sub>[m]</sub>, smooth<sub>[m]</sub> stuff*; *during* in the phrase *during this time, this someone's fingers<sub>[m]</sub> move as this someone wants*; *bit* in the phrase *because of this, a little bit of something like hot<sub>[m]</sub> water<sub>[m]</sub> moves*; *way* in the phrase *sometimes when someone is drinking<sub>[m]</sub> something in this way*; *as* in the phrase *part of the edge<sub>[m]</sub> at the top<sub>[m]</sub> of this thing touches one of this someone's lips<sub>[m]</sub> for a short time, as this someone wants*. According to Cliff Goddard, most of these cases are allolexes of different primes. The word *stuff* should be replaced with the expression *the same something*, while *during* could be seen as an allolex of the expression *at this time*, used in relation to time periods (personal discussion November 2011). The word *bit* was used in some earlier explications, instead, it should have been replaced with the relatively recently-proposed prime LITTLE. The word *way* is an allolex of the prime THIS. Similar questions arise from the explication of *God* (Wierzbicka 2001: 21; see Appendix 2), as words other than primes or molecules occur there as well (*exist, always, they (them)*). Solutions like these are undoubtedly understandable and even desirable, but they should be clearly marked, reasoned – and opened to wider use. Various small and

---

<sup>20</sup> Molecules are mini explications that function as units in the semantic structure of other, yet more complex words.

unimportant adjustments may be needed even in the explications made by the real experts. The same kind of flexible attitude could be shown with more visible and frequent things, as suggested above in relation to the English allolexes of the prime SOMEONE. It is understandable, though, that any changes in the current prime inventory would require major consideration and extensive exploration.

At the present point in time, one may be justified in asking whether it is even possible to think about any ultimate pure NSM explications if all the NSM versions are making them with slightly different toolboxes. We would not bother too much about this question. What we do think, however, is that the current set of the NSM primes is, after the decades of careful description, identifiable in a relatively language independent way, and that this set can be regarded as the pure core of the NSM.

As soon as the English-based NSM version gets any kind priority among NSM versions, significant risks arise at two directions: First, favoring underlying lexico-grammatical features typical only to the English language would lead to Anglocentrism. On the other side, restricting the English version of NSM from creating sufficient allolexy would lead to too complicated explications in the most popular version of NSM.

From one important aspect, the English version of NSM is very special compared to the others. Namely, it is through the English-based NSM that most linguists (and other potentially interested people) form their attitude to the whole approach. Adding appropriate allolexes and grammatical features to the English variant would simply make the English-based NSM explications easier to comprehend by a wider audience. The English-based NSM should be treated as any other version of the NSM, it is not the universal NSM concept itself. To make it concrete, we could add another column to Table 1 in section 2, and separate the universal NSM primes and the English based NSM variants into different columns, the latter one having more and flexible allolexes.

Taken together, the main theoretical message of this paper is that the difference between the underlying ultimate universal NSM concepts and the English-based NSM version should be kept clear. This requirement may sound theoretical, unachievable, and even unnecessary, but this is the final point to which all of the practical question marks and concerns raised in the present study have led us. The findings in this paper pointed out some quite vague underlying constructions within the general system of primes. Regardless of the reported shortcomings, the NSM has shown indisputable

ability to deal with an extremely challenging task, namely semantic analysis.

## References

- Amberber, Mengistu (2008) Semantic primes in Amharic. In Cliff Goddard (ed.), pp. 83–120. Amsterdam: John Benjamins.
- Bugenhagen, Robert D. (2002) The syntax of semantic primitives in Mangaaba-Mbula. In Cliff Goddard & Anna Wierzbicka (eds.), Vol. 2. pp. 1–64. Amsterdam/Philadelphia: John Benjamins.
- Chappell, Hilary (1994) The universal syntax of semantic primes in Mandarin Chinese. In Cliff Goddard & Anna Wierzbicka (eds.), Vol. 1, pp. 243–322. Amsterdam/Philadelphia: John Benjamins.
- Diller, Anthony (1994) Thai. In Cliff Goddard & Anna Wierzbicka (eds.), pp. 149–170. Amsterdam: John Benjamins.
- van Driem, George (2004) Review: Cliff Goddard & Anna Wierzbicka (eds.), *Meaning and universal grammar: Theory and empirical findings*. 2 vols. *Language* 80 (1): 163–165.
- Enfield, Nick J. (2002) Combinatoric properties of Natural Semantic Metalanguage expressions in Lao. In Cliff Goddard & Anna Wierzbicka (eds.), Vol. 2. pp. 145–256. Amsterdam/Philadelphia: John Benjamins.
- Geeraerts, Dirk (2010) *Theories of Lexical Semantics*. Oxford University Press.
- Gladkova, Anna (2007) *Russian Emotions, Attitudes and Values: Selected Topics in Cultural Semantics*. PhD Thesis. Australian National University.
- (2010) *Russian cultural semantics: Emotions, values, attitudes*. Moscow: Languages of Slavonic Cultures. (in Russian)
- Goddard, Cliff (2002) Semantic primes and universal grammar in Malay (Bahasa Melayu). In Cliff Goddard & Anna Wierzbicka (eds.), Vol. 1, pp. 87–172. Amsterdam/Philadelphia: John Benjamins.
- (ed.) (2008), *Cross-Linguistic Semantics*. Studies in Language Companion Series 102. Amsterdam: John Benjamins.
- (2011a) *Semantic Analysis: A Practical Introduction*, second, revised edition. Oxford: Oxford University Press.
- (2011b) The lexical semantics of *language* (with special reference to *words*). *Language Sciences* 33: 40–57.
- (2011c) Chart of NSM semantic primes <[http://www.griffith.edu.au/\\_\\_data/assets/pdf\\_file/0003/347565/nsm-chart-01-2011.pdf](http://www.griffith.edu.au/__data/assets/pdf_file/0003/347565/nsm-chart-01-2011.pdf)> (4 February 2013).
- (2014) Have to, Have got to, and Must: NSM analyses of English modal verbs of “necessity”. In Maite Taboada & Radoslava Trnavac (eds.), *Nonveridicality and Evaluation: Theoretical, computational and corpus approaches*, pp. 50–75. Leiden: Brill.
- Goddard, Cliff & Karlsson, Susanna (2008) Re-thinking THINK in contrastive perspective: Swedish vs. English. In Cliff Goddard (ed.), pp. 225–240. Amsterdam: John Benjamins.

- Goddard, Cliff & Wierzbicka, Anna (eds.) (1994) *Semantic and Lexical Universals*. Studies in Language Companion Series 25. Amsterdam: John Benjamins.
- (eds.) (2002) *Meaning and Universal Grammar: Theory and Empirical Findings*, Vol. 1–2. Studies in Language Companion Series. Amsterdam/Philadelphia: John Benjamins.
- Habib, Sandy (2011) Ghosts, fairies, elves, and nymphs: Towards a semantic template for non-human being concepts. *Australian Journal of Linguistics* 31: 411–443.
- Junker, Marie-Odile (2008) Semantic primes and their grammar in a polysynthetic language: East Cree. In Cliff Goddard (ed.), pp. 163–204. Amsterdam: John Benjamins.
- McCawley, James D. (1983) Review of Anna Wierzbicka's *Lingua Mentalis: The semantics of natural language*. *Language* 59 (3): 654–659.
- Murray, D.W & Button, Gregory (1988) Human emotions: Some problems of Wierzbicka's 'Simples'. *American Anthropologist* 90 (3): 684–686.
- Nirvi, Ruben E. (1936) Kansankielen synonyymeistä ja synonyymisluonteisista sanoista. *Virittäjä* 40: 24–88.
- Onishi, Masayuki (1994) Semantic primitives in Japanese. In Cliff Goddard & Anna Wierzbicka (eds.), pp. 361–386. Amsterdam/Philadelphia: John Benjamins.
- Oxford English Dictionary <<http://www.oed.com>> (29 September 2014)
- Peeters, Bert (1994) Semantic and lexical universals in French. In Cliff Goddard & Anna Wierzbicka (eds.), pp. 423–442. Amsterdam: John Benjamins.
- Sinnemäki, Kaius (2011) *Language Universals and Linguistic Complexity*. PhD thesis. General Linguistics, Department of Modern Languages, University of Helsinki.
- The NSM home page: <<http://www.griffith.edu.au/humanities-languages/school-languages-linguistics/research/natural-semantic-metalanguage-homepage>>
- Travis, Catherine (2002) *La Megalengua Semántica Natural: The Natural Semantic Metalanguage of Spanish*. In Cliff Goddard & Anna Wierzbicka (eds.), Vol. 1, pp. 173–242. Amsterdam/Philadelphia: John Benjamins.
- Tuovila, Seija (2005) *Kun on tunteet: Suomen kielen tunnesanojen semantiikkaa* [Such emotions]. The semantics of emotion words in the Finnish language] Doctoral dissertation. Oulu: Oulun yliopisto.
- Yoon, Kyung-Joo (2008) The Natural Semantic Metalanguage of Korean. In Cliff Goddard (ed.), pp. 121–162. Amsterdam: John Benjamins.
- (2006) *Constructing a Korean Natural Semantic Metalanguage*. Seoul: Hankook Publishing Co.
- Vanhatalo, Ulla (forthcoming): *Things you can do with the Natural Semantic Metalanguage*.
- Vanhatalo, Ulla & Tissari, Heli (forthcoming): *Natural Semantic Metalanguagen suomenkielinen versio 1.0* [Natural Semantic Metalanguage. The Finnish version 1.0].
- Wawrzyniak, Jan K. (2010) Native speakers, mother tongues and natural semantic metalanguages. *Language Sciences* 32: 648–670.
- Wierzbicka, Anna (1972) *Semantic Primitives*. Frankfurt: Athenäum.
- (1988) Semantic Primitives: A Rejoinder to Murray and Button. *American Anthropologist, New Series* 90 (3): 686–689.

- (1992) *Semantics, Culture and Cognition: Universal Human Concepts in Culture-Specific Configurations*. New York: Oxford University Press.
- (1994) Semantic primitives across languages: A Critical review. In Cliff Goddard & Anna Wierzbicka (eds.), pp. 445–500. Amsterdam: John Benjamins.
- (1996) *Semantics: Primes and Universals*. Oxford: Oxford University Press.
- (2001) *What Did Jesus Mean? Explaining the Sermon on the Mount and the Parables in Simple and Universal Human Concepts*. New York: Oxford University Press.
- (2002) Semantic primes and universal grammar in Polish. In Cliff Goddard & Anna Wierzbicka (eds.), pp. 65–144. Amsterdam/Philadelphia: John Benjamins.
- (2004) The English expressions *good boy* and *good girl* and cultural models of child rearing. *Culture & Psychology* 10(3): 251–278.

Contact information:

Ulla Vanhatalo  
School of Language, Translation and Literary Studies  
33014 University of Tampere

The Department of Modern Languages  
PO Box 24  
00014 University of Helsinki  
Finland  
ulla(dot)vanhatalo(at)helsinki(dot)fi

Heli Tissari  
University of Eastern Finland  
Philosophical Faculty  
Joensuu Campus  
School of Humanities  
PO Box 111  
80101 Joensuu  
Finland  
heli(dot)tissari(at)uef(dot)fi

Anna Idström  
University of Helsinki  
anna(dot)idstrom(at)welho(dot)com

**Appendix 1. Some examples of the basic combinatorial possibilities of NSM primes (from Goddard 2011c). For the full set in Finnish, see Vanhatalo & Tissari, forthcoming.**

**MINÄ (I), SINÄ (YOU)**

*minä en tiedä* ‘I don’t know’

*minä haluan sinun tekevän/tietävän jotakin* ‘I want you to do/know something’

*jotain pahaa voi tapahtua minulle/sinulle* ‘something bad can happen to me/you’

*joku kuten minä/minunlaiseni* ‘someone like me’

**PALJON~MONI (MUCH~MANY)**

*paljon ihmisiä / monet ihmiset* ‘many people’

*paljon asioita / monet asiat* ‘many things’

*paljon osia / monet osat* ‘many parts’

*monenlaisia* ‘many kinds’

*monina aikoina* ‘at many times’

*monissa paikoissa* ‘in many places’

*paljon jotain tällaista (esim. vettä)* ‘much something of this kind (e.g. water)’

*paljon enemmän* ‘much/many more’

**PAHA (BAD)**

*jotain pahaa* ‘something bad’

*pahat ihmiset* ‘bad people’

*jotain pahaa tapahtuu* ‘something bad happens’

*tehdä jotain pahaa (jollekulle)* ‘do something bad (to someone)’

*tuntea jotain pahaa* ‘feel something bad’

*tämä on paha(a)* ‘this is bad’

*on paha jos...* ‘it is bad if...’

**TAHTOA (WANT)**

*joku tahtoo jotakin* ‘someone wants something’

*joku tahtoo tehdä/tietää/sanoa jotakin* ‘someone wants to do/know/say something’

*joku tahtoo jonkun muun tekevän/tietävän jotakin* ‘someone wants someone else to do/know something’

*joku tahtoo jotakin tapahtuvan* ‘someone wants something to happen’

## Appendix 2. Some examples of the Finnish based NSM in use.

### Discourse particle *well* (Engl.)

*Well*, —

(Goddard 2011a:173)

<p>you said something a short time before because of this I want to say something after a very short time I'm thinking about it now because I want to say it well [I say: — ]</p>	<p>sinä sanoit jotain vähän aikaa sitten tämän takia minä tahdon sanoa jotain hyvin lyhyen ajan kuluttua minä ajattelen sitä nyt koska minä tahdon sanoa sen hyvin [minä sanon: —]</p>
---	--

### Noun *God* (Engl.)

*God*

(Wierzbicka 2001: 21)

<p>(a) God is someone (not something) (b) this someone is someone good (c) this someone is not someone like people (d) there isn't anyone else like this someone (e) this someone exists always (f) everything exists because this someone wants it to exist (g) people exist because this someone wants them to exist (h) this someone exists because this someone exists, not because of anything else (i) this someone lives</p>	<p>(a) Jumala on joku (ei jokin) (b) hän on joku hyvä (c) hän ei ole sellainen kuin ihmiset (d) ei ole ketään muuta sellaista kuin hän (e) hän on olemassa aina (f) kaikki on olemassa koska hän tahtoo sen olevan olemassa (g) ihmiset ovat olemassa koska hän tahtoo heidän olevan olemassa (h) hän on olemassa koska hän on olemassa, ei minkään muun takia (i) hän elää</p>
---	---

**Mari Wiklund**

**La transmission des effets stylistiques des phrases sans verbe  
fini dans les traductions finnoises –  
L'exemple des pièces de théâtre de Jean-Paul Sartre**

**Résumé**

Cette étude portera sur la traduction des phrases sans verbe fini dans les versions finnoises des *Mains sales* (1948) et de *Huis clos* (1947) de Jean-Paul Sartre. La construction sans verbe fini est très fréquente dans les pièces de théâtre où le texte consiste principalement en des dialogues. Normalement, c'est surtout le verbe de la phrase qui sert à l'ancrer dans une situation (Riegel, Pellat & Rioul 2004). Par conséquent, en raison de l'absence du verbe, l'interprétation de ce type de phrases dépend du contexte. Le phénomène met en valeur le caractère propre des pièces de théâtre ; le fait que les dialogues soient destinés à être joués sur scène où l'ancrage situationnel est assuré par de nombreux moyens multimodaux. En créant un effet d'oralité et d'expressivité, le phénomène sert à la dramatisation de la parole. Il constitue aussi un procédé de mise en relief. Dans les traductions finnoises, la construction originale n'est cependant pas toujours conservée ; un tiers des phrases sans verbe fini ont été traduites sans conserver la construction de la phrase source. Le problème qui se pose consiste à savoir si les effets stylistiques véhiculés par la phrase source se transmettent dans la traduction lorsque la structure syntaxique n'y est pas conservée.

**1. Introduction**

Comme de nombreuses études l'ont déjà montré, les phrases sans verbe fini constituent une ressource stylistique essentielle dans la presse écrite et dans la littérature française (Drillon 1991 ; Lefeuvre 1999, 2007 ; Noailly 2002 ; Larjavaara 2003 ; Riegel, Pellat & Rioul 2004 ; Havu 2009, 2010 ; Combettes & Kuyumcuyan 2010 ; Lehtinen 2011). Par 'une phrase sans verbe fini', nous entendons une séquence qui ne comporte pas de verbe conjugué et qui est entourée de deux signes de ponctuation forte (le point, le point d'exclamation, le point d'interrogation, les points de suspension).

Cette séquence, ainsi que la séquence suivante, commence par une majuscule (Larjavaara 2003). Nous avons choisi d'employer ici le terme de 'phrase sans verbe fini' au lieu de 'phrase averbale', puisque Lefeuvre (1999 : 19) compte parmi les phrases averbales uniquement « les énoncés qui comportent de façon assurée un prédicat averbal ». Ainsi, elle ne prend pas compte par exemple d'énoncés construits autour d'un infinitif ou d'une interjection. Comme nous avons choisi de tenir compte de tous les énoncés qui ne comportent pas de verbe fini (aussi des infinitifs et des interjections), nous préférons utiliser ici le terme de 'phrase sans verbe fini' qui a un sens plus large. Riegel, Pellat et Rioul (2004), quant à eux, utilisent le terme de 'phrase nominale'. Nous n'employons pas ce terme non plus, puisque, à l'instar de Lefeuvre (1999 : 19), le terme 'nominal' nous semble un peu trop restrictif.

Les phrases sans verbe fini sont souvent considérées comme étant proches du langage parlé (Gardes-Tamine 1990 : 26 ; Lefeuvre 1999 : 84). Par conséquent, elles peuvent servir notamment à créer un effet d'oralité et d'expressivité dans un texte (Larjavaara 2003 ; Tuomarla 2004). Tuomarla (2004 : 332) remarque que dans la presse écrite, l'émiettement du texte résultant de l'usage de phrases courtes dépourvues du verbe simule le rythme du langage oral. Selon Drillon (1991 : 131), si ces phrases courtes se terminent par un point, celui-ci leur donne un 'caractère affirmatif'.

Selon Riegel, Pellat et Rioul (2004 : 457) les phrases sans verbe fini sont particulièrement fréquentes dans les phrases exclamatives, mais la construction est tout à fait possible aussi dans les phrases déclaratives, interrogatives et impératives. Normalement, c'est surtout le verbe de la phrase qui sert à l'ancrer dans une situation (Riegel, Pellat & Rioul 2004 : 457). Par conséquent, une phrase sans verbe diffère d'une phrase verbale en ce qui concerne l'ancrage situationnel : une phrase sans verbe est « avant tout plus sensible aux variations de la situation d'énonciation particulière » et « elle manifeste souvent une plus grande expressivité que la phrase canonique » (Riegel, Pellat & Rioul 2004 : 457). De même, selon Larjavaara (2003 : 62–63), c'est le verbe fini de la phrase qui fait progresser le texte ; les phrases sans verbe fini, quant à elles, servent plutôt à 'fixer le regard' du lecteur ou à mettre en relief la séquence concernée. Ce type de construction peut aussi constituer un point de transition permettant par exemple un changement de point de vue (Noailly 2002 : 144). Parfois, il sert tout simplement à préciser ou à expliquer ce qui vient d'être dit (Larjavaara 2003 : 62). Évidemment, les différentes fonctions stylistiques des phrases sans verbe fini ne s'excluent pas ; la même occurrence peut par

exemple servir à mettre en relief la séquence concernée, à créer un effet d'oralité et à donner un caractère affirmatif à la phrase en question.

Cette étude sera centrée sur les différents effets stylistiques véhiculés par les phrases sans verbe fini et encore plus particulièrement, sur la façon de laquelle ces effets sont transmis dans leurs traductions finnoises.<sup>1</sup> Le corpus étudié consiste en deux pièces de théâtre : *Les Mains sales* (Sartre 1948) et *Huis clos* (Sartre 1947) ainsi que les traductions finnoises de ces ouvrages, *Likaiset kädet* (Kaukonen 1966) et *Suljetut ovet* (Rankkala 1966). Les analyses seront basées sur des données quantitatives classifiant les différentes stratégies utilisées par les traducteurs pour interpréter en finnois la construction syntaxique du texte source.<sup>2</sup> Les deux traductions datent de 1966. Cela joue un certain rôle sur le plan lexical. Autrement dit, les choix lexicaux des traducteurs ne correspondent pas toujours à ceux d'un traducteur travaillant en 2014. Néanmoins, comme la syntaxe d'une langue change beaucoup plus lentement que le lexique, cet aspect temporel ne semble pas être pertinent du point de vue structurel (Costaouec 2009). Par conséquent, il n'a pas d'influence sur les stratégies de traduction analysées dans cet article.

Les phrases sans verbe fini existent aussi en finnois (Helasvuo 1997 ; Hakulinen, Vilkkuna & Korhonen 2004 : 839–840, 974). Il s'agit le plus souvent des syntagmes nominaux ou adjectivaux au nominatif ou au partitif (Hakulinen, Vilkkuna & Korhonen 2004 : 974). On trouve cependant aussi des syntagmes post- ou prépositionnels ainsi que des infinitifs et des participes (*ibid.*). Helasvuo (1997) a consacré une partie de sa thèse aux phrases nominales « libres » dans un corpus du finnois parlé. La phrase nominale a été considérée comme 'libre' lorsqu'elle n'a pas pu être analysée comme étant un constituant syntaxique d'une autre proposition (Helasvuo 1997 : 127). Selon l'auteur (p. 141), ces phrases nominales libres (*free NPs*) peuvent avoir quatre types de fonctions discursives : 1) Elles peuvent servir à *identifier* un référent, une proposition ou une séquence plus longue (p. 142). 2) Parfois, elles *classifient* une entité en attribuant un nom à la classe à laquelle cette entité appartient (*ibid.*). 3) Lorsqu'une phrase nominale au locatif donne un prédicat à une phrase nominale au nominatif, il s'agit d'une *construction avec un thème et une*

<sup>1</sup> Nous remercions sincèrement les relecteurs de cet article. Leurs commentaires pertinents nous ont permis d'améliorer le texte considérablement.

<sup>2</sup> Nous avons déjà abordé la traduction finnoise des phrases sans verbe fini dans *Les Mains sales* de Jean-Paul Sartre dans un petit article publié dans les actes d'un colloque (Lehtinen 2011).

*orientation* (p. 150). La phrase nominale au nominatif y constitue le thème, et la phrase nominale au locatif y constitue l'orientation. A la différence des phrases nominales qui servent à identifier ou à classifier (*cf.* 1–2), cette construction nominale commence quelque chose de nouveau dans le discours. 4) Les phrases nominales libres peuvent également constituer des *constructions introduisant un topique*. Dans ce cas, elles sont interprétées à la lumière de ce qui va suivre, et elles projettent une continuation d'échanges sur le topique qui est introduit (p. 153–154). Les trois premières fonctions sont prédicatives, tandis que la dernière est d'ordre référentiel (p. 141). En conclusion, Helasvuo (1997 : 141) constate que les fonctions discursives des phrases nominales libres du finnois semblent être les mêmes que celles découvertes par Ono et Thompson (1994) dans un corpus anglais.

Havu (2009) a étudié les occurrences des phrases sans verbe fini dans un corpus de romans d'auteurs finnois (Hotakainen, Härkönen, Joensuu, Lardot), parus dans les années 1990 et 2000. Elle remarque que l'emploi de phrases sans verbe fini est intimement lié à l'auteur et au style du roman. On peut trouver plus de phrases sans verbe dans certains romans avec peu de dialogues que dans d'autres comprenant de nombreux dialogues (Havu 2009 : 207). Dans les pièces de théâtre dont le texte consiste principalement en des dialogues, leur fréquence semble malgré tout être très élevée. Cela est certainement lié au fait que la relation entre les dialogues des pièces de théâtre et le langage parlé est très concrète ; il ne s'agit pas seulement d'une simulation fictive, mais les dialogues sont vraiment destinés à être présentés oralement.

Havu (2010) a également étudié la traduction de ce type de phrases dans un corpus de romans d'auteurs français (Duras, Echenoz, Gavalda) traduits en finnois. Elle a remarqué que les prédications averbales étaient très fréquentes dans l'œuvre de Gavalda, bien représentées chez Duras et peu nombreuses dans le roman d'Echenoz (Havu 2010). En ce qui concerne la traduction finnoise de ces phrases, il y avait aussi des différences notables : le traducteur d'Echenoz a transformé la plupart des constructions averbales en phrases verbales, tandis que les traducteurs de Duras et de Gavalda ont le plus souvent conservé la structure de la phrase source (Havu 2010).

La traduction finnoise des phrases sans verbe fini n'a cependant pas encore été étudiée dans les pièces de théâtre ni dans l'œuvre de Sartre. Néanmoins, en raison de leur relation proche avec le langage oral, les pièces de théâtre nous semblent particulièrement intéressantes sur ce plan.

Effectivement, Sartre fait grand usage des phrases sans verbe fini surtout dans ses pièces de théâtre. L’ancrage situationnel et le rôle du verbe y étant assuré par d’autres moyens – tels que les descriptions détaillées du cadre spatio-temporel et des réactions des personnages – l’interprétation des phrases est possible bien qu’elles ne comportent pas de verbe (Riegel, Pellat & Rioul 2004 : 458). Dans les versions finnoises des *Mains sales* et de *Huis clos* un tiers des phrases sans verbe fini sont cependant traduites sans conserver la construction syntaxique originale. Dans ces cas, ne pas perdre les effets stylistiques véhiculés par la phrase source constitue un défi pour les traducteurs.

Les phrases sans verbe fini sont caractéristiques des dialogues des pièces de théâtre étudiées. Les didascalies qui donnent des indications à la régie et aux acteurs y prennent également bien souvent la forme averbale. Naturellement, les didascalies ne véhiculent pas les mêmes effets stylistiques que les phrases sans verbe fini apparaissant dans les répliques. Elles ont cependant été retenues dans cette étude, puisque d’une manière générale, les didascalies semblent être concernées par les mêmes stratégies de traduction que les autres phrases sans verbe fini. De plus, dans un certain nombre de cas, les traducteurs changent la construction syntaxique d’une didascalie, et cela a pour conséquence la modification du cadre interprétatif de la phrase.

## 2. Différentes stratégies de traduction des phrases sans verbe fini

Nous commençons par une brève présentation de données quantitatives concernant les différentes stratégies de traduction des phrases sans verbe fini.

La version originale des *Mains sales* (Sartre 1948) comporte en tout 1198 phrases sans verbe fini. L’ouvrage comprend 247 pages. Il y a donc en moyenne 4,9 phrases sans verbe fini par page. La version originale de *Huis clos* (Sartre 1947), quant à lui, comprend 95 pages. Le nombre de phrases sans verbe fini y est de 593, ce qui correspond à 6,2 occurrences par page. Dans la traduction finnoise des *Mains sales* (*Likaiset kädet*, 1966), la construction syntaxique originale a été conservée dans 67,9 % des cas. En ce qui concerne *Huis clos* (*Suljetut ovet*, 1966), le pourcentage correspondant est de 68,3 %. Il est notable que le pourcentage est pratiquement le même dans les deux œuvres traduites, bien que le traducteur ne soit pas le même. Le fait que la structure de la phrase soit changée dans un tiers des cas s’explique probablement avant tout par des

raisons stylistiques liées à ce qui est propre au finnois écrit (Newmark 1988 ; Gambier 2008). En effet, notre hypothèse est que les phrases sans verbe fini constituent un trait plus marqué dans les textes finnois que dans les textes français. Pour confirmer cette hypothèse, il faudrait cependant dépouiller un grand corpus de textes différents ; comme Havu (2010) l'a constaté, certains traducteurs ont tendance à transformer les constructions averbales en phrases verbales, tandis que d'autres ont plutôt tendance à les conserver. D'un autre côté, le nombre moins important des phrases sans verbe fini dans les versions finnoises peut s'expliquer aussi par une tendance des traducteurs à une écriture texto-centrique au lieu d'une écriture scéno-centrique, caractérisée notamment par des éléments déictiques tels que les phrases sans verbe fini.

Dans les deux ouvrages, il semble y avoir trois stratégies différentes pour traduire les phrases sans verbe fini sans conserver la construction originale. Le changement de la structure syntaxique de la phrase source constitue évidemment une stratégie de traduction bien connue (Chesterman 1997 ; Gambier 2008). Le but de cette partie de notre étude consiste à décrire plus en détail les manières selon lesquelles le changement de la structure syntaxique est effectué dans le corpus étudié. Les stratégies les plus fréquentes consistent à ajouter un verbe fini à la phrase en question ou bien, à la relier à la phrase précédente ou à la phrase suivante (*cf.* aussi Havu 2010). Il y a aussi un certain nombre de cas où la phrase sans verbe fini apparaissant dans le texte français a été complètement supprimée par le traducteur. Cela n'est cependant pas étonnant, puisque l'omission est un phénomène courant dans la traduction (Vázquez Ayora 1977 ; Chesterman 1997 ; Molina & Hurtado Albir 2002 ; Pym 2010). Les stratégies mentionnées seront illustrées plus bas à l'aide d'exemples.

L'ordre de fréquences de ces différentes stratégies est le même dans les deux ouvrages, mais le traducteur des *Mains sales* (*Likaiset kädet*, 1966) recourt à la stratégie consistant à relier les phrases (12,7 %) presque aussi souvent qu'à la stratégie consistant à ajouter un verbe à la phrase (13,9 %). Dans la traduction de *Huis clos* (*Suljetut ovet*, 1966), en revanche, les cas où un verbe a été ajouté (17,0 %) sont beaucoup plus fréquents que les cas où deux ou plusieurs phrases ont été reliées (7,8 %). Les nombres d'occurrences et les pourcentages correspondants de ces différentes stratégies de traduction sont présentés dans le tableau 1.

**Tableau 1** : Fréquences des différentes stratégies de traduction dans le corpus étudié

STRATÉGIE DE TRADUCTION	<i>Les mains sales</i> Nombre d'occurrences (%)	<i>Huis clos</i> Nombre d'occurrences (%)
Conserver la construction	814 (67,9 %)	405 (68,3 %)
Ajouter un verbe	167 (13,9 %)	101 (17,0 %)
Relier des phrases	152 (12,7 %)	46 (7,8 %)
Phrase supprimée	65 (5,4 %)	41 (6,9 %)

La classification obtenue par ces analyses quantitatives servira de base pour les analyses qualitatives présentées ci-dessous. Nous nous y focaliserons sur les trois stratégies dans lesquelles la construction originale n'est pas conservée, mais au passage, nous nous interrogeons aussi sur les cas où la construction sans verbe fini n'a pas été modifiée. Nous nous poserons la question de savoir si le changement éventuel de la structure syntaxique joue un rôle sur les effets stylistiques véhiculés par la phrase sans verbe fini apparaissant dans le texte source. Si la structure n'a pas été changée, nous nous interrogeons sur la question de savoir si les effets stylistiques qu'elle véhicule sont les mêmes dans les deux langues.

## 2.1 La première stratégie : ajouter un verbe

Dans la version finnoise des *Mains sales* (*Likaiset kädet*, 1966), un verbe a été ajouté à 13,9 % des phrases sans verbe fini apparaissant dans le texte source. L'exemple (1) présente un extrait où cette stratégie est employée trois fois.

- (1) TEXTE ORIGINAL (EN FRANÇAIS) :
- 01 *Les mêmes, Olga.*
  - 02 *Elle pose une valise près de la porte.*
  - 03 → *Olga (à Ivan) : Voilà. Tu pourras la fixer sur ton porte-bagages ?*
  - 04 → *Ivan : Montre. Oui. Très bien.*
  - 05 *Olga : Il est dix heures. Tu peux filer. On t'a dit pour le barrage et la maison.*
  - 06 *Ivan : Oui.*
  - 07 *Olga : Alors bonne chance.*
  - 08 *Ivan : Parle pas de malheur. (Un temps.) Tu m'embrasses ?*
  - 09 → *Olga : Bien sûr.*
  - 10 *Elle l'embrasse sur les joues.*
- (Les mains sales. Sartre 1948 : 41).

## (1) TRADUCTION (EN FINNOIS) :

01 (*Samat ja Olga, joka asettaa matkalaukun ovenpieleen.*)

02 → *Olga (Ivanille): Siinä se on. Voitko kulettää sen pakettitelineellä?*

03 → *Ivan: Näytähän. Kyllä se menee.*

04 *Olga: Kello on kaksikymmentäkaksi. Voit lähteä. Tulosi tiedetään padolla ja 05 talossa.*

06 *Ivan: Niin.*

07 *Olga: Onnea matkalle.*

08 *Ivan: Kunhan ei kova onni. (tauko) Etkö suutele minua?*

09 → *Olga: Tietenkin suutelen. (suutelee Ivania molemmille poskille)*

(Likaiset kädet. Sartre 1948 : 22. Traduit par Toini Kaukonen en 1966).

L'extrait du texte source comprend huit phrases sans verbe fini. Quatre de ces phrases ont été traduites avec une phrase contenant un verbe fini. Dans le texte original, la première réplique d'Olga commence par 'Voilà.' (ligne 03). L'interprétation de cette séquence est complètement dépendante de la description de l'action d'Olga donnée juste avant sa réplique : *Elle pose une valise près de la porte* (ligne 02). Dans la version finnoise, le verbe 'être' (*olla* ; PRÉ, 3<sup>e</sup> P. SG.) est ajouté à cette séquence. Elle est traduite 'Siinä se on.' (ligne 02), ce qui correspond à 'là' (*siinä*) + 'elle' (*se*) + 'est' (*on*) → 'elle est là'. Le changement de la construction syntaxique n'est pas obligatoire dans ce cas. En effet, dans ce contexte, il serait tout à fait possible de traduire 'Voilà.' par exemple par '*Kas tässä.*'<sup>3</sup> ou par '*Tässä.*' qui ne comportent pas de verbe fini et dont les interprétations seraient plus dépendantes du contexte que celle de '*Siinä se on.*'. Le fait d'ajouter un verbe à cette phrase sert à l'ancrer dans la situation (Riegel, Pellat & Rioul 2004 : 457).<sup>4</sup> Autrement dit, le traducteur passe ici d'un repérage déictique situationnel, qui n'est pas verbalisé, à un repérage anaphorique.<sup>5</sup> De même, la traduction finnoise ressemble moins au langage oral que la séquence originale, puisque la phrase '*Siinä se on.*' est plus longue que 'Voilà.' Ainsi, elle efface le rythme que la phrase courte crée dans la réplique originale. 'Voilà.' pourrait également être traduit ici par exemple par '*Tässä.*' qui serait aussi court que 'Voilà.' et qui conserverait ainsi le rythme de la réplique originale.

<sup>3</sup> *Kas tässä* est une des traductions finnoises typiquement données pour le mot 'voilà' dans les dictionnaires bilingues (Kalmbach & Sundelin 2000 [2008] : 1174).

<sup>4</sup> La remarque de Riegel, Pellat et Rioul (2004 : 457) est basée sur un corpus français. Au moins dans ce cas précis, elle semble cependant s'appliquer au finnois aussi.

<sup>5</sup> Selon Rabatel (2000), les présentatifs du français – tels que 'voilà' – cumulent des repérages déictiques et pseudo-anaphoriques.

Dans la réplique suivante d'Ivan (ligne 04), il y a deux phrases sans verbe fini – ‘Oui.’ et ‘Très bien.’ – qui sont traduites en une seule phrase ‘*Kyllä se menee.*’ (ligne 03). Dans ce cas, c’est le verbe ‘aller’ (*mennä* ; PRÉ, 3<sup>e</sup> P. SG.) qui est ajouté : ‘d’accord’ (*kyllä*) + ‘elle’ (*se*) + ‘va’ (*menee*) → ‘d’accord, ça va aller’. Le fait d’ajouter un verbe n’est pas obligatoire ici non plus. Le choix du traducteur est probablement motivé par des facteurs pragmatiques et stylistiques. Les séquences sans verbe fini constituent ici la réponse à la question d’Olga : ‘Tu pourras la fixer sur ton porte-bagages ?’ (ligne 03). *Kyllä* (‘d’accord’) est utilisé pour annuler l’hésitation qu’exprime la première partie de la réponse d’Ivan (‘Montre.’, ligne 04) (Hakulinen *et al.*, 2004 : 1150). Il serait possible de répondre à ce type d’interrogation totale uniquement par *kyllä*, mais en raison de la brièveté de la réponse, la réponse pourrait alors être considérée comme un peu rude.

Une traduction littérale conservant la structure sans verbe fini (‘*Kyllä. Oikein hyvin.*’) serait possible aussi, mais les faits d’ajouter un verbe et de relier les deux séquences en une seule unité rendent la réponse plus polie et plus idiomatique (Newmark 1988 ; Hakulinen *et al.* 2004 : 1150). Cela est mis en valeur par le fait que la réplique originale consiste en fait en trois phrases courtes : ‘Montre. Oui. Très bien.’ (ligne 04). En finnois, l’émiettement de la réplique en trois parties si courtes (‘*Näytähän. Kyllä. Oikein hyvin.*’) semblerait un peu heurté sur le plan stylistique. Le problème qui se pose, c’est que la réplique originale (ligne 04) évoque fortement le rythme de l’oral, et le changement de la structure syntaxique effectué par le traducteur a pour conséquence la dilution de cet effet. Cela signifie, ici comme plus haut, un changement de perspective. Autrement dit, Sartre adopte ici une écriture *scéno-centrique*, déictique, mimant les interactions orales au plus près. Le traducteur, quant à lui, adopte une version plus écrite, de nature *texto-centrique*.

La troisième occurrence apparaît à la fin de l’extrait, dans la dernière réplique d’Olga : ‘Bien sûr.’ (ligne 09). Cette réplique est traduite en reprenant le verbe ‘embrasser’ (*suudella* ; PRÉ, 1<sup>re</sup> P. SG.), qui apparaît dans la réplique précédente d’Ivan : ‘*Tietenkin suutelen.*’ (ligne 09), ‘bien sûr’ (*tietenkin*) + ‘j’embrasse’ (*suutelen*) → ‘bien sûr que je t’embrasse’. La traduction correspond bien à l’usage réel du finnois dans une situation de conversation spontanée. En effet, il s’agit ici d’une paire adjacente question-réponse (Sacks, Schegloff & Jefferson 1974 ; Traverso 1999), et il est tout à fait typique que le verbe d’une interrogation totale soit repris dans la seconde partie de paire (Hakulinen *et al.* 2004 : 1148–1149). Le fait que

le verbe soit répété ici souligne la réponse positive d'Olga. Sur le plan grammatical, la réponse sans verbe formulée uniquement avec le mot *tietenkin* ('bien sûr'), serait correcte aussi, mais elle serait facilement considérée comme impolie dans ce type de contexte. En effet, la variable *politesse* nous semble capitale ici, indépendamment du changement de perspective scéno-centrique ou texto-centrique. À la différence des traductions des occurrences précédentes, ce changement de la construction syntaxique est pratiquement obligatoire ici.

Cette occurrence est similaire au cas de 'Voilà.' (ligne 03) dans le sens où ici aussi, l'interprétation de la phrase source est complètement dépendante du contexte, tandis que la traduction finnoise l'est beaucoup moins grâce à la présence du verbe. D'un côté, ce type de répliques sans verbe met en valeur le caractère propre des pièces de théâtre ; le fait que les dialogues y sont destinés à être joués sur scène où l'ancrage situationnel des répliques est assuré par de nombreux moyens multimodaux (cadre, action, mimiques et gestes, direction du regard, musique, lumière, décors, costumes, etc.). De ce point de vue, il est dommage que la structure sans verbe fini ne soit pas conservée. Néanmoins, comme déjà dit, dans ce cas précis, le fait qu'un verbe soit ajouté à la réplique correspond à l'usage idiomatique.

L'extrait inclut aussi trois occurrences qui sont traduites sans changer la construction syntaxique : La réplique courte d'Ivan, 'Oui.' (ligne 06), est traduite '*Niin.*' ('Oui'). La réplique d'Olga, 'Alors bonne chance.' (ligne 07) est traduite '*Onnea matkalle.*' ('Bonne chance'), et la didascalie 'Un temps.' (entre parenthèses, ligne 08) est traduite '*tauko*' ('un temps'), sans signe de ponctuation et sans majuscule. Dans tous ces trois cas, le fait de conserver la construction sans verbe fini est la seule possibilité de traduction. Il ne serait pas logique d'y ajouter un verbe, et les autres stratégies de traduction mentionnées dans le tableau 1 n'y seraient pas possibles non plus. Dans les deux premières occurrences (lignes 06 et 07), la construction sans verbe fini est employée pour créer un effet d'oralité et pour évoquer un dialogue spontané authentique. Le style d'écriture *scéno-centrique* adopté par Sartre est donc conservé ici. Ainsi, les effets stylistiques du texte source sont transmis à la traduction finnoise. En ce qui concerne la didascalie 'Un temps.' (ligne 08), l'emploi d'une construction sans verbe est la seule possibilité, et la construction n'y véhicule pas d'effets stylistiques particuliers. (À strictement parler, la construction n'est pas complètement conservée ici, puisque la traduction '*tauko*' ne commence pas par une majuscule et elle ne se termine pas par un point. De

ce fait, dans la traduction finnoise, la didascalie n'a pas de statut de phrase. Sur le plan stylistique, cela n'a cependant pas d'importance dans ce contexte.)

Le deuxième exemple présente un autre extrait comportant des occurrences de ce type. Cet extrait a aussi été pris dans *Les Mains sales*.

(2) TEXTE ORIGINAL (EN FRANÇAIS) :

01 Hoederer : *Ça te donne envie de rire quand on t'embrasse ? (Jessica baissa la tête.)*

03 → *Hein ?*

04 Jessica : *Oui.*

05 Hoederer : *Alors, tu es froide ?*

06 Jessica : *C'est ce qu'ils disent.*

07 Hoederer : *Et toi, qu'en penses-tu ?*

08 Jessica : *Je ne sais pas.*

09 → Hoederer : *Voyons. (Il l'embrasse.) Eh bien ?*

10 Jessica : *Ça ne m'a pas donné envie de rire.*

11 *La porte s'ouvre, Hugo entre.*

(Les mains sales. Sartre 1948 : 227–228).

(2) TRADUCTION (EN FINNOIS) :

01 Hoederer: *Vai naurattaa sinua, kun sinua suudellaan? (Jessica painaa häpeillen päänsä alas.)*

03 → *Niinhän se oli?*

04 Jessica: *Niin.*

05 Hoederer: *Sinä olet siis kylmä?*

06 Jessica: *Niin sanotaan.*

07 Hoederer: *Entä mitä itse luulet?*

08 Jessica: *En tiedä.*

09 → Hoederer: *Saammepä nähdä. (suutelee häntä) No, miltä tuntui?*

10 Jessica: *Minua ei naurattanut. (ovi aukenee, Hugo astuu sisään)*

(Likaiset kädet. Sartre 1948 : 126. Traduit par Toini Kaukonen en 1966).

Dans l'extrait français, il y a trois phrases sans verbe fini : 'Hein ?' (ligne 03), 'Oui.' (ligne 04) et 'Eh bien ?' (ligne 09). Seulement la deuxième, 'Oui.' (ligne 04), est traduite en finnois sans changer la construction syntaxique ('Niin.', 'Oui.'). Comme dans les occurrences de l'exemple 1 traduites sans changer la construction syntaxique, ici aussi, le fait de conserver la construction est la seule stratégie de traduction possible. Dans les deux langues, cette réplique consistant en une seule particule discursive évoque l'interaction orale spontanée. L'auteur et le traducteur adoptent ici une perspective scéno-centrique, caractérisée notamment par ce type de

répliques, dont l'interprétation est complètement dépendante de leur contexte d'occurrence.

Dans les deux autres cas, un verbe est ajouté, bien que le changement de la construction syntaxique n'y soit pas obligatoire. 'Hein ?' (ligne 03) est traduit '*Niinhän se oli?*' (ligne 03) en ajoutant le verbe 'être' (*olla* ; IMP, 3<sup>e</sup> P. SG.) : 'comme ça' + ENC (*niinhän*) + 'ce' (*se*) + 'était' (*oli*) → 'c'était bien comme ça ?'. Il s'agit d'une question de vérification d'Hoederer ajoutée à la fin de son propre tour de parole (Hakulinen *et al.* 2004 : 1156). La présence du verbe ne serait pas obligatoire dans la traduction, puisque dans ce contexte, il serait tout à fait possible de traduire 'hein' par exemple par la particule *niinkö* (*niin* + enclitique interrogatif -*kO*) (Hakulinen *et al.* 2004 : 773). Sur le plan pragmatique, *niinkö* marcherait très bien ici, et son usage est fréquent dans le langage courant. De plus, il est moins long et moins soigné que '*Niinhän se oli?*'. Par conséquent, il conserverait peut être mieux l'effet d'oralité véhiculé par le 'hein' du français. Comme il ne semble pas y avoir d'autre raison pour la transformation de la structure, ce choix est probablement motivé par la volonté du traducteur de limiter le nombre de phrases sans verbe fini ainsi que par ses conceptions de la norme en générale et de celle du texte théâtral. Il est possible qu'il s'agisse d'une façon de faire qui lui est habituelle, qui fait qu'il ne respecte pas les spécificités du style oral.

De même, 'Eh bien ?' (ligne 09) est traduit '*No, miltä tuntui?*', avec le verbe 'sentir' (*tuntea* ; IMP, 3<sup>e</sup> P. SG.) : 'eh bien' (*no*) + 'comment' (*miltä*) + 'sentait' (*tuntui*) → 'eh bien, comment tu t'es sentie ?'. Ici aussi, 'eh bien' pourrait être traduit uniquement par la particule *no* (Kalmbach & Sundelin 2000 [2008] : 374 ; Hakulinen *et al.* 2004 : 999). Le traducteur a cependant opté pour une formulation plus explicite. L'avantage de cette traduction est qu'elle est complètement privée d'ambiguïtés. La particule *no* suivie d'un point d'interrogation serait suffisante aussi, puisque la description de l'action qui la précède donne assez d'informations pour son interprétation. Néanmoins, il est vrai que par rapport à la formulation explicite, de nature texto-centrique ('*No, miltä tuntui?*'), son interprétation nécessiterait plus d'effort de la part du lecteur. Peut-être paradoxalement, le défaut de la traduction explicite est, lui aussi, lié à cette facilité d'interprétation. En effet, l'explicitation diminue la nécessité du lecteur de se mettre à la situation fictive, d'imaginer les événements de la pièce tels qu'ils se dérouleraient sur scène. Cette traduction ne conserve donc pas la perspective scéno-centrique adoptée par l'auteur du texte source.

Le troisième exemple a été tiré de *Huis clos*. Dans la traduction finnoise de cet ouvrage, la stratégie consistant à ajouter un verbe a été employée dans 17,0 % des cas. L'extrait comporte huit occurrences de ce type.

(3) TEXTE ORIGINAL (EN FRANÇAIS) :

- 01 Inès : *Le bourreau, c'est chacun de nous pour les deux autres.*  
 02 *Un temps. Ils digèrent la nouvelle.*  
 03 *Garcin (d'une voix douce) : Je ne serai pas votre bourreau. Je ne vous veux*  
 04 → *aucun mal et je n'ai rien à faire avec vous. Rien. C'est tout à fait simple.*  
 05 → *Alors voilà : chacun dans son coin ; c'est la parade. Vous ici, vous ici, moi*  
 06 → *là. Et du silence. Pas un mot : ce n'est pas difficile, n'est-ce pas ? Chacun*  
 07 *de nous a assez à faire avec lui-même. Je crois que je pourrais rester dix*  
 08 *mille ans sans parler.*  
 09 *Estelle : Il faut que je me taise ?*  
 10 → *Garcin : Oui. Et nous... nous serons sauvés. Se taire. Regarder en soi, ne*  
 11 → *jamais lever la tête. C'est d'accord ?*  
 12 → *Inès : D'accord.*  
 13 → *Estelle (après hésitation) : D'accord.*  
 (Huis clos. Sartre 1947 : 42–43).

(3) TRADUCTION (EN FINNOIS) :

- 01 *Ines: Jokainen meistä on kahden toisen pyöveli.*  
 02 *(Tauko. Kaikki pohtivat tätä uutista)*  
 03 *Garcin (hiljaa): Minä en aio olla teidän pyövelinne. Minä en tahdo teille*  
 04 *mitään pahaa enkä halua olla teidän kanssanne missään tekemisissä.*  
 05 → *En tahdo mitään. Se on aivan yksinkertaista. Jäämme tänne jokainen*  
 06 → *nurkkaamme, vältämme toisiamme. Te olette siellä, te siellä ja minä*  
 07 → *täällä. Olemme aivan hiljaa. Emme sano sanaakaan, eihän se ole*  
 08 *vaikeaa, vai mitä? Jokaisella meistä on aivan tarpeeksi tekemistä omassa*  
 09 *itsessämme. Luulen, että voisin olla puhumatta kymmenentuhatta vuotta.*  
 10 *Estelle: Pitääkö minun olla hiljaa?*  
 11 → *Garcin: Pitää. Silloin... silloin pelastumme. Pitää olla vaiti. Pitää katsoa*  
 12 *vain itseensä, ei nostaa koskaan päätään. Suostutteko?*  
 13 → *Ines: Suostun.*  
 14 → *Estelle (epäröiden hetken): Suostun.*  
 (Suljetut ovet. Sartre 1947 : 154. Traduit par Marja Rankkala en 1966).

La première réplique de Garcin (lignes 03–08) comporte trois phrases sans verbe fini. Toutes ces phrases ont été traduites avec un verbe. La première occurrence, ‘Rien.’ (ligne 04), est traduite en ajoutant le verbe ‘vouloir’ (*tahtoa* ; NÉG, PRÉ, 1<sup>re</sup> P. SG.) à la forme négative → ‘*En tahdo mitään.*’, ‘je ne veux rien’. Le fait d’ajouter le verbe ‘vouloir’ n’est pas obligatoire

ici. En fait, dans ce cas, la traduction n'est même pas tout à fait correcte, puisque 'Rien.' se réfère clairement à la proposition précédente, 'je n'ai rien à faire avec vous' (ligne 04). En répétant le mot 'rien', la phrase sans verbe fini met en valeur le contenu de la proposition précédente. La traduction ne transmet pas vraiment cet effet de mise en valeur, puisque la phrase '*En tahdo mitään.*', 'je ne veux rien', ajoute plutôt une information par rapport à ce qui a été dit : Garcin ne veut aucun mal à Inès et à Estelle et il ne veut avoir rien à faire avec elles, mais il ne veut pas autre chose non plus. A notre avis, la traduction '*en missään*' aurait été plus adéquate ici, puisque dans la phrase précédente de la version finnoise (ligne 04), Garcin dit '*enkä halua olla teidän kanssanne missään tekemisissä*' ('et je ne veux avoir rien à faire avec vous'). En reprenant l'équivalent finnois de 'rien' dans ce contexte, cette traduction aurait conservé l'effet de mise en valeur de la phrase source.

De même, la phrase originale ('Rien.') est très courte ; la traduction est nettement plus longue. Par conséquent, le rythme créé par la traduction n'est pas le même. En effet, selon Chafe (1988 : 397), les lecteurs éprouvent une image auditive du texte qu'ils lisent. L'image auditive de la phrase 'Rien.' évoque l'oralité grâce à sa brièveté, mais il n'en va pas de même pour '*En tahdo mitään.*', qui semble stylistiquement plus neutre. Toujours en raison de sa brièveté, la phrase originale est très expressive, et elle arrête le regard du lecteur d'une manière efficace (Larjavaara 2003). La traduction atténue ces effets stylistiques. Autrement dit, elle donne un sens particulier à ce qui a une portée plus générale ; elle donne à un propos philosophique de portée générale une dimension factuelle. L'interprétation de la présence de la dimension factuelle est possible dans le texte original aussi, mais celle-ci s'y cumule avec la dimension philosophique. Aussi bien sur le plan des idées que sur le plan formel, cette traduction appauvrit considérablement la dramatisation et l'expressivité véhiculées par la phrase sans verbe fini. De ce fait, on peut constater que la traduction n'est pas adéquate sur le plan sémantico-énonciatif.

Ensuite, 'Vous ici, vous ici, moi là.' (lignes 05–06) ainsi que la phrase suivante, 'Et du silence.' (ligne 06) sont traduites avec le verbe 'être' (*olla* ; PRÉ, 2<sup>e</sup> P. PL.) : '*Te olette siellä, te siellä ja minä täällä.*' ('vous êtes là, vous là et moi ici.', lignes 06–07), et (*olla* ; PRÉ, 1<sup>re</sup> P. PL.) '*Olemme aivan hiljaa*' ('nous sommes sans rien dire', ligne 07). Ces traductions expriment explicitement le verbe 'être' qui est implicitement présent dans les phrases originales. La structure sans verbe fini aurait pu être conservée dans les deux cas. Il serait tout à fait possible de traduire 'Vous ici, vous ici, moi là.'

sans ajouter le verbe ‘être’ (*‘Te siellä, te siellä ja minä täällä.’*), et ‘Et du silence.’ pourrait être traduit par exemple par *‘Ja sitten hiljaa.’* En ajoutant le verbe, le traducteur facilite le travail interprétatif du lecteur. Ces phrases ne seraient cependant pas ambiguës même sans verbe, puisque la présence implicite du verbe ‘être’ y est tellement évidente. De plus, *‘Olemme aivan hiljaa’* est immédiatement suivi par *‘Emme sano sanaakaan,’* (‘nous ne disons pas un mot’, ligne 07). En fait, en facilitant le travail interprétatif du lecteur, ce type de traductions où un verbe est ajouté bien que le changement averbal ne soit pas obligatoire appauvrit considérablement et inutilement le texte source.

De plus, comme dans le cas de ‘Rien.’ (ligne 04), ici aussi, le fait d’ajouter un verbe rallonge les phrases. Cela a pour conséquence un changement du rythme : dans le texte original, la réplique de Garcin comporte une suite de séquences courtes qui créent un effet d’un rythme saccadé à l’oral (lignes 04–06). Il s’agit de la dramatisation de la parole et de la mise en valeur de la relation interhumaine. Dans la version finnoise, ces effets sont atténués. Par conséquent, la perspective scéno-centrique du texte source n’y est pas conservée.

La réplique suivante de Garcin (lignes 10–11) comporte, elle aussi, trois phrases sans verbe fini. Dans la traduction finnoise, il y a un verbe fini dans toutes ces phrases. ‘Oui’ (ligne 10), au début de la réplique, est traduit en reprenant le verbe ‘ falloir’ (*pitää*, ZÉRO, PRÉ) qui apparaît dans la réplique précédente : ‘Il faut que je me taise ?’ (ligne 09), *‘Pitääkö minun olla hiljaa?’* (ligne 10); ‘Oui’ → *‘Pitää.’* (ligne 11 ; ‘il le faut’). La traduction correspond à l’usage réel de la langue finnoise dans ce type de situations, puisque la réponse minimale à une interrogation totale y consiste typiquement en la répétition du verbe de la question, lorsqu’il s’agit d’une ‘vraie question recherchant une information’ (Hakulinen *et al.* 2004 : 1147–1148). Ensuite, la discussion continue d’une manière ou d’une autre à traiter le même sujet, comme c’est le cas dans cet exemple aussi (*ibid.*). ‘Kyllä’ (‘oui’) ne serait pas une traduction impossible ici, mais dans ce cas, la question d’Estelle (ligne 10) serait considérée plutôt comme une question de vérification et non pas comme une question qui recherche vraiment une information (*ibid.*). Le fait de reprendre le verbe de la question met en valeur, d’un côté, l’incertitude d’Estelle et de l’autre côté, la fermeté de Garcin.

Le même verbe (*pitää*, ZÉRO, PRÉ) est utilisé dans les phrases suivantes : ‘Se taire.’ (ligne 10) est traduit *‘Pitää olla vaiti.’* (ligne 11 ; ‘il faut se taire.’), et ‘Regarder en soi, ne jamais lever la tête.’ (lignes 10–11)

est traduit '*Pitää katsoa vain itseensä, ei nostaa koskaan päätään.*' (lignes 11–12 ; 'il faut regarder en soi, ne jamais lever la tête.'). Dans ces deux cas, l'emploi d'un verbe fini est la seule possibilité, puisque – à la différence du français – l'infinitif du finnois ne peut pas remplacer l'impératif. Le fait de répéter le verbe *pitää* trois fois dans la même réplique ne serait cependant pas obligatoire ; il y aurait d'autres verbes qui pourraient être utilisés à sa place. Néanmoins, cette répétition contribue à la création d'un ton insistant, qui est présent dans le texte original aussi, mais qui y est créé par d'autres moyens (succession de séquences courtes, absence du verbe, etc.). Le choix d'une modalité déontique effectué ici est cependant discutable. Il serait possible de dire 'vouloir' / 'ne pas faire' autrement qu'en restant au constat. Ici aussi, une pluralité de sens est possible. Le contexte fait interpréter, mais même en contexte, plusieurs interprétations restent possibles selon la signification globale que l'on se fie des relations humaines et de la philosophie sartrienne sous-jacente.

De même, les répliques d'Inès et d'Estelle (lignes 12–13) sont traduites en reprenant le verbe 'accepter' (*suostua*) apparaissant à la fin de la réplique de Garcin dans la version finnoise. 'C'est d'accord ?' (ligne 11) est traduit '*Suostutteko?*' (ligne 12 ; 'vous l'acceptez ?' / 'vous êtes d'accord ?'), et le même verbe (*suostua*, PRÉ, 1<sup>re</sup> P. SG.) est employé pour traduire 'D'accord.' (lignes 12–13) dans les répliques d'Inès et d'Estelle (lignes 13–14 ; '*Suostun.*', 'je l'accepte.'). Le changement de la structure syntaxique n'est pas grammaticalement imposé ici, mais comme la question de Garcin (ligne 12) est clairement une vraie question de recherche d'information, le fait que le verbe soit repris dans les réponses correspond à l'usage réel et semble stylistiquement adapté à ce contexte (Hakulinen *et al.* 2004 : 1147–1148).

En principe, la répétition du verbe de la question pourrait être remplacée par la particule *kyllä* ('oui'), qui peut aussi constituer une réponse minimale à une interrogation totale (*ibid.*). Néanmoins, ici encore plus que dans le cas du verbe *pitää* repris au début de la réplique de Garcin (ligne 11), la répétition du verbe met en valeur le fait qu'il ne s'agit pas seulement d'une question de vérification mais d'une vraie question à laquelle la réponse peut être soit positive soit négative. Dans ce type de contextes, le fait de ne pas avoir conservé la construction sans verbe fini semble donc complètement justifié.

Cet exemple inclut une seule occurrence de phrase sans verbe fini traduite sans changer la construction syntaxique. Il s'agit de la didascalie 'Un temps' (ligne 02) qui est traduite '*Tauko.*' ('Un temps'). Dans ce cas,

la construction n'a pas de fonctions stylistiques, et le changement de la forme de phrase ne serait pas possible.

## 2.2 La deuxième stratégie : relier à une autre phrase

Dans la version finnoise des *Mains sales* (*Likaiset kädet*, 1966), 12,7 % des phrases sans verbe fini sont traduites en les reliant à la phrase précédente ou suivante. Dans la traduction de *Huis clos* (*Suljetut ovet*, 1966), le pourcentage correspondant est de 7,8 %. La séquence à laquelle la phrase sans verbe fini est reliée peut être soit une phrase verbale, soit une phrase sans verbe. Dans le premier cas, il s'agit donc d'une manière d'ajouter un verbe à la phrase qui n'en avait pas à l'origine. Dans le deuxième cas, la construction sans verbe est conservée, mais le nombre total de phrases de ce type diminue.

L'exemple (4) présente un cas où une phrase sans verbe fini est reliée à la phrase suivante qui est une phrase verbale. L'extrait a été pris dans *Les Mains sales*. Dans la traduction de cet ouvrage, la stratégie consistant à relier les phrases a été employée dans 12,7 % des cas.

### (4) TEXTE ORIGINAL (EN FRANÇAIS) :

- 01 Hugo : *Qui parle de guerre civile ? Hoederer, je ne vous comprends pas ; il*  
 02 *suffirait d'un peu de patience. Vous l'avez dit vous-même : l'Armée*  
 03 *rouge chassera le Régent et nous aurons le pouvoir pour nous seuls.*  
 04 Hoederer : *Et comment ferons-nous pour le garder ? (Un temps.) Quand*  
 05 *l'Armée rouge aura franchi nos frontières, je te garantis qu'il y aura de*  
 06 *durs moments à passer.*  
 07 Hugo : *L'Armée rouge...*  
 08 → Hoederer : *Oui, oui. Je sais. Moi aussi, je l'attends. Et avec impatience.*  
 09 *Mais il faut bien que tu le dises : toutes les armées en guerre,*  
 10 *libératrices ou non, se ressemblent : elles vivent sur le pays*  
 11 *occupé. [...]*

(Les mains sales. Sartre 1948 : 192).

## (4) TRADUCTION (EN FINNOIS) :

01 *Hugo: Kuka tässä puhuu kansalaissodasta? Hoederer, minä en ymmärrä teitä.*

02 *Ei tarvita muuta kuin hiukan kärsivällisyyttä. Itsekin sanoitte, että puna-*

03 *armeija karkoittaa sijaishallitsijan ja me pääsemme valtaan.*

04 *Hoederer: Entä miten me säilyttäisimme tämän vallan käsissämme?*

05 *(tauko) Voit olla varma, että meille tulee vaikeat oltavat heti kun*

06 *puna-armeija on ylittänyt rajamme.*

07 *Hugo: Puna-armeija...*

08 → *Hoederer: Niin, niin, minä tiedän ja minä odotan, minäkin. Vieläpä varsin*

09 *malttamattomasti. Mutta sinun on muistettava, että kaikki armeijat,*

10 *olivatpa sitten vapauttajia tai eivät, muistuttavat toisiaan sodassa,*

11 *elämällä valloitetun maan kustannuksella. [...]*

(Likaiset kädet. Sartre 1948 : 105. Traduit par Toini Kaukonen en 1966).

L'extrait français comprend quatre phrases sans verbe fini. Dans trois cas, la construction syntaxique est conservée : 'Un temps.' (entre parenthèses, ligne 04) est traduit '*tauko*' (entre parenthèses, ligne 05) ; 'L'Armée rouge...' (ligne 07) est traduit '*Puna-armeija...*' (ligne 07) ; et 'Et avec impatience.' (ligne 08) est traduit '*Vieläpä varsin malttamattomasti.*' (lignes 08–09). Dans les deux premiers cas, le fait de conserver la construction syntaxique est la seule possibilité. La didascalie 'Un temps' ne véhicule pas d'effets stylistiques particuliers qui seraient liés à sa construction syntaxique. La réplique consistant en le groupe nominal 'L'Armée rouge...' suivi de points de suspension (ligne 07) est plus intéressante. Ici, la construction sans verbe fini sert à dramatiser la parole, et les points de suspension accentuent cet effet. En effet, selon Catach (1996 : 63), les points de suspension « expriment l'inaccompli » et « ils rejoignent, d'une certaine façon, le non-dit, mais un non-dit explicite, expressif ». Ces effets stylistiques sont transmis à la version finnoise où la forme de la phrase ainsi que les points de suspension sont conservés. La réplique en question constitue aussi une phrase nominale libre (Helasvuo 1997). Plus précisément, il s'agit d'une construction introduisant un topique (pp. 153–154) : Hugo reprend l'élément 'L'Armée rouge...' ('*Puna-armeija...*', ligne 07) de la réplique précédente d'Hoederer (lignes 02–03) et l'introduit comme un topique. Ce topique va être traité dans la réplique suivante d'Hoederer (lignes 08–11).

La troisième occurrence, 'Et avec impatience.' (ligne 08), quant à elle, aurait, en principe, pu être reliée à la phrase précédente. Le traducteur a cependant choisi de conserver la forme de la phrase. Les effets de mise en valeur et de dramatisation véhiculés par la phrase source sont donc transmis à la version finnoise. D'un autre côté, dans la traduction finnoise, cette

occurrence est précédée d'une phrase relativement longue reliant trois phrases courtes du texte source. Il aurait été difficile d'y relier encore une quatrième phrase. Par conséquent, en raison des choix antérieurs du traducteur, le fait de conserver la structure syntaxique est pratiquement obligatoire ici aussi.

Le quatrième cas ('Oui. Oui.', ligne 08), quant à lui, est relié à la phrase suivante par une virgule. Cette phrase ('Je sais.', '*minä tiedän*') comporte un verbe fini ('savoir', *tietää* ; PRÉ, 1<sup>re</sup> P. SG.). La séquence qui se forme ainsi est ensuite associée à la phrase suivante par la conjonction de coordination *ja* ('et'). Par conséquent, la séquence 'Oui, oui. Je sais. Moi aussi, je l'attends.' (ligne 08) du texte original devient dans la traduction finnoise '*Niin, niin, minä tiedän ja minä odotan, minäkin.*' (ligne 08), 'oui, oui, je sais et j'attends, moi aussi'. Le fait d'avoir choisi la particule *niin* au lieu de *kyllä* est justifié, puisque *niin* constitue la réponse minimale typique de questions focalisées portant sur un autre élément que le verbe de la phrase (Hakulinen *et al.* 2004 : 1152). Ici, il s'agit d'une réaction à la réplique d'Hugo (ligne 07), qui consiste donc uniquement en un groupe nominal aussi bien dans le texte original ('L'Armée rouge...') que dans la traduction ('Puna-armeija...').

La traduction '*Niin, niin. Minä tiedän. Minäkin odotan sitä.*' où la structure du texte original est conservée serait, en principe, tout à fait possible aussi. Le changement de la structure de la phrase n'est donc pas obligatoire ici. Le choix du traducteur évite cependant de fragmenter la réplique avec plusieurs phrases sans verbe fini. En effet, comme déjà dit, tout de suite après les phrases reliées (ligne 08), il y a une autre phrase sans verbe fini ('Et avec impatience.'). De la même manière que l'ajout d'un verbe, le fait de relier les phrases a pour conséquence un changement du rythme. Dans le texte original, le début de la réplique d'Hoederer (ligne 08) comporte une suite de séquences courtes imitant un certain rythme de l'oral. Dans ce contexte, ce rythme crée l'effet d'un ton insistant. Dans la version finnoise, le fait de relier les phrases atténue cet effet stylistique. La perspective scéno-centrique du texte source est donc remplacée par la perspective texto-centrique dans la traduction finnoise.

L'exemple (1) présente un cas de ce même type. Les phrases 'Les mêmes, Olga. Elle pose une valise près de la porte.' (lignes 01–02) y sont traduites en une seule phrase en les reliant par le pronom relatif *joka* ('qui') : '*Samat ja Olga, joka asettaa matkalaukun ovenpieleen.*' (ligne 01) → 'les mêmes et Olga qui pose une valise près de la porte'. Le changement de la structure syntaxique ne serait pas obligatoire dans ce cas. Il est

cependant notable que le verbe apparaît ici seulement dans la proposition subordonnée ; la proposition principale ne comporte pas de verbe. Par conséquent, cette expression pourrait en fait être considérée comme une phrase sans verbe fini. De toute façon, ce qui est intéressant ici sur le plan stylistique, c'est que, pour une fois, il ne s'agit pas de la création d'un effet d'oralité dans un dialogue, mais il s'agit d'une didascalie fournissant une description du cadre de la scène. Dans le texte français, les personnages de la situation sont indiqués sur une ligne, et l'action d'Olga sur la ligne suivante. Cela met en valeur le déroulement de la scène d'une pièce de théâtre : d'abord, les spectateurs voient les personnages sur scène et ensuite, il y a l'action d'Olga. La structure du texte original met en valeur ces deux phases, tandis que la traduction finnoise les présente parallèlement.

L'exemple (5), quant à lui, illustre une occurrence typique d'un cas où deux séquences dont ni l'une ni l'autre ne comporte un verbe sont reliées. L'extrait a été pris dans *Huis clos* où 7,8 % des phrases sans verbe fini ont été traduites en les reliant à une autre phrase.

(5) TEXTE ORIGINAL (EN FRANÇAIS) :

01 Inès : *Vous avez beaucoup souffert ?*

02 Estelle : *Non. J'étais plutôt abrutie.*

03 Inès : *Qu'est-ce que... ?*

04 Estelle : *Une pneumonie. (Même jeu que précédemment.) Eh bien, ça y*

05 → *est, ils s'en vont. Bonjour ! Bonjour ! Que de poignées de*

06 *main. Mon mari est malade de chagrin, il est resté à la maison.*

07 (A Inès.) *Et vous ?*

08 Inès : *Le gaz.*

(Huis clos. Sartre 1947 : 30–31).

(5) TRADUCTION (EN FINNOIS) :

01 Ines: *Saitteko kärsiä paljon?*

02 Estelle: *En. Minä vain heikkenin vähitellen.*

03 Ines: *Miten siis?*

04 Estelle: *Keuhkokuume. (jatkaa samaan tapaan kuin edellä) No niin, nyt ne*

05 → *lähtevät tiehensä. Hyvää päivää, hyvää päivää. Siinäpä puristetaan*

06 *kättä. Mieheni on surusta sairas, hän jää kotiin. (Inekselle) Entä te?*

07 Ines: *Kaasua.*

(Suljetut ovet. Sartre 1947 : 147–148. Traduit par Marja Rankkala en 1966).

L'extrait du texte source comporte neuf phrases sans verbe fini : 'Non.' (ligne 02), 'Une pneumonie' (ligne 04), 'Même jeu que précédemment.' (entre parenthèses, ligne 04), 'Bonjour !' (ligne 05), 'Bonjour !' (ligne 05),

‘Que de poignées de main.’ (lignes 05–06), ‘A Inès.’ (entre parenthèses, ligne 07), ‘Et vous ?’ (ligne 07) et ‘Le gaz.’ (ligne 08). Dans quatre cas, la construction sans verbe est conservée. Trois séquences sont traduites en ajoutant un verbe fini, et les deux qui restent sont reliées l’une à l’autre par une virgule. Il s’agit des phrases exclamatives ‘Bonjour ! Bonjour !’ (ligne 05). Dans la version finnoise, ces phrases ne sont pas seulement reliées l’une à l’autre, mais elles sont aussi transformées en une phrase déclarative en remplaçant le point d’exclamation final par un point : ‘*Hyvää päivää, hyvää päivää.*’ (ligne 05). En effet, les phrases exclamatives sans verbe fini sont très souvent transformées en phrases déclaratives dans le corpus étudié. Naturellement, cela diminue leur niveau d’expressivité. Sur le plan grammatical, la séquence pourrait être transposée telle quelle en la traduisant ‘*Päivää ! Päivää !*’. Selon nous, l’emploi de deux points d’exclamation consécutifs serait cependant plus marqué dans un texte finnois que dans un texte français. En finnois, les points d’exclamation constitueraient des indices d’une vraie exclamation plutôt que des indices d’expressivité. ‘*Päivää, päivää !*’, avec un seul point d’exclamation serait plus neutre, mais dans ce contexte, il ne serait pas logique sur le plan rythmique. ‘*Hyvää päivää, hyvää päivää.*’ constitue un choix neutre et cohérent. Néanmoins, cette traduction est beaucoup moins expressive que ‘Bonjour ! Bonjour !’ qui arrête vraiment le regard du lecteur (Larjavaara 2003).

Les quatre occurrences traduites sans changer la structure syntaxique sont : ‘Une pneumonie.’ (ligne 04), ‘A Inès.’ (entre parenthèses, ligne 07), ‘Et vous ?’ (ligne 07) et ‘Le gaz.’ (ligne 08). Dans les cas des phrases ‘A Inès.’ et ‘Et vous ?’, le fait de conserver la structure syntaxique est la seule possibilité. En ce qui concerne les phrases ‘Une pneumonie.’ et ‘Le gaz.’, le traducteur aurait pu y ajouter un verbe. Cela aurait facilité le travail interprétatif du lecteur, mais en même temps, les effets stylistiques véhiculés par ces phrases sans verbe fini auraient été perdus. Il s’agit de phrases nominales libres qui répondent à la question posée par Inès (ligne 03) et qui servent à identifier un référent (Helasvuo 1997 : 142). Cette fonction discursive est typique du langage oral (Ono & Thompson 1994 ; Helasvuo 1997). Par conséquent, ces occurrences véhiculent un effet d’oralité qui disparaîtrait si la construction syntaxique était changée.

L’exemple suivant présente un autre cas de ce même type où deux ou plusieurs phrases sans verbe fini sont reliées. L’extrait a été tiré des *Mains sales*.

## (6) TEXTE ORIGINAL (EN FRANÇAIS) :

01 *Georges et Slick restent hésitants sur le pas de la porte.*

02 *Hoederer : Eh bien ? Qu'est-ce que vous attendez ? Vous avez compris ?*

03 *Slick : On croyait...*

04 *Hoederer : Il n'y a rien à croire, faites ce qu'on vous dit.*

05 → *Slick : Bon. Bon. Bon.*

06 *Georges : C'était pas la peine de faire toutes ces histoires.*

(Les mains sales. Sartre 1948 : 100).

## (6) TRADUCTION (EN FINNOIS) :

01 (*Georges ja Slick seisovat empien ovella.*)

02 *Hoederer: Mitä te siinä odotatte? Ymmärsittekö?*

03 *Slick: Me luultiin...*

04 *Hoederer: Ei tässä ole mitään luulemista. Tehkää niin kuin sanotaan.*

05 → *Slick: Oolrait, oolrait.*

06 *Georges: Maksoks vaivaa pitää tämmöstä melua tyhjästä?*

(Likaiset kädet. Sartre 1948 : 52–53. Traduit par Toini Kaukonen en 1966).

Dans cet exemple, il y a quatre phrases sans verbe fini. La première, ‘Eh bien ?’ (ligne 02), n’est pas incluse dans la traduction finnoise. De même, dans le texte source, la deuxième réplique de Slick consiste en trois particules discursives dont chacun constitue une phrase sans verbe fini : ‘Bon. Bon. Bon.’ (ligne 05). Sur le plan grammatical, il serait tout à fait possible de transposer cette réplique à la version finnoise telle quelle, sans changer la construction syntaxique. Néanmoins, dans la traduction finnoise, le nombre des particules est réduit à deux, et elles sont reliées en une seule phrase par une virgule : ‘*Oolrait, oolrait.*’ (ligne 05). La traduction correspond bien à l’usage réel, puisque les directives à impératif sont typiquement reçues avec une particule en finnois (Hakulinen *et al.* 2004 : 1161–1162). La réduction du nombre des particules est certainement liée au fait que le mot argotique finnois *oolrait* est nettement plus long que le mot ‘bon’ du texte original. En raison des longueurs des mots et de la ponctuation différentes, le rythme de la traduction ‘*Oolrait, oolrait.*’ ne correspond pas à celui de la réplique originale ‘Bon. Bon. Bon.’. Le rythme de cette dernière est plus saccadé. Néanmoins, selon nous, le fait de diviser la réplique en trois petites séquences par la présence d’un point après chaque particule ne serait pas vraiment propre au finnois ; le choix du traducteur nous semble stylistiquement plus adapté (Newmark 1988 ; Molina & Hurtado 2002).

L’extrait ne comporte pas d’occurrences de phrases sans verbe fini qui seraient traduites sans changer la construction syntaxique.

### 2.3 La troisième stratégie : supprimer la phrase

Comme l'exemple (6) ci-dessus l'a déjà montré, il arrive aussi qu'une phrase sans verbe fini soit complètement supprimée. Cela est évidemment tout à fait courant dans la traduction (Vázquez Ayora 1977 ; Chesterman 1997 ; Molina & Hurtado Albir 2002 ; Pym 2010).<sup>6</sup> L'exemple (7) illustre aussi un cas de ce type. L'extrait a été tiré des *Mains sales* où la fréquence de cette stratégie est de 5,4 %.

(7) TEXTE ORIGINAL (EN FRANÇAIS) :

01 *Olga va chercher une assiette, du pain et du jambon. Pendant qu'elle dispose*  
 02 *l'assiette et les aliments sur la table, devant lui, il parle :*  
 03 → *Hugo : Je ne me suis pas trompé, pour ta chambre. Pas une fois. Tout est*  
 04 *comme dans mon souvenir. (Un temps.) Seulement quand j'étais*  
 05 *en taule, je me disais : c'est un souvenir. La vraie chambre est là-bas,*  
 06 *de l'autre côté du mur. Je suis entré, j'ai regardé ta chambre et elle*  
 07 *n'avait pas l'air plus vraie que mon souvenir. La cellule aussi, c'était*  
 08 *un rêve. Et les yeux d'Hoederer, le jour où j'ai tiré sur lui. Tu crois*  
 09 *que j'ai une chance de me réveiller ? Peut-être quand tes copains*  
 10 *viendront sur moi avec leurs joujoux...*  
 (Les mains sales. Sartre 1948 : 31).

(7) TRADUCTION (EN FINNOIS) :

01 (*Olga ottaa kaapista kinkkua ja leipää ja asettaa lautaset Hugon eteen*  
 02 *sillävälän kun tämä jo puhuu*)  
 03 → *Hugo:*<sup>7</sup> *En ollut erehtynyt huoneestasi. Kaikki on aivan kuin muistin. (tauko)*  
 04 *Vankilassa sanoin itselleni, ettei se ole kuin pelkkä muisto ja että*  
 05 *o i k e a huone on muurin toisella puolen. Kun astuin äsken sisään,*  
 06 *katsoin ympärilleni ja havaitsin, ettei huone ollut muistikuvaani*  
 07 *todempi. Sellikin on unta. Ja Hoedererin silmät sellaisina kuin ne näin*  
 08 *ampuessani silloin. Luuletko, että minä vielä herään? Ehkä silloin, kun*  
 09 *toverisi käyvät kimppuuni leluineen.*  
 (Likaiset kädet. Sartre 1948 : 17. Traduit par Toini Kaukonen en 1966).

Dans cet exemple, une phrase sans verbe fini ('Pas une fois.', ligne 03) a été complètement supprimée par le traducteur. La séquence 'Je ne me suis pas trompé, pour ta chambre. Pas une fois. Tout est comme dans mon

<sup>6</sup> Bien qu'il s'agisse d'une stratégie de traduction bien courante, cela ne signifie pas que le fait de supprimer une séquence soit toujours bien justifié.

<sup>7</sup> Le nom du personnage qui parle ('Hugo') n'a pas été mentionné dans la traduction finnoise. Comme il est clair qu'il s'agit d'une coquille, nous nous sommes permise de l'ajouter.

souvenir.’ (lignes 03–04) est traduite : ‘*En ollut erehtynyt huoneestasi. Kaikki on aivan kuin muistin.*’ (ligne 02), ‘je ne me suis pas trompé sur ta chambre ; tout est exactement comme je me souvenais’. Le fait de supprimer la phrase sans verbe fini n’est pas obligatoire ici ; elle pourrait être traduite sans changer la construction syntaxique ou bien, elle pourrait être reliée à la phrase précédente. Dans ce contexte, il est évident que la phrase sans verbe fini véhicule une fonction de mise en valeur : elle sert à souligner le contenu de la première phrase de la réplique, ‘Je ne me suis pas trompé, pour ta chambre.’. La traduction finnoise, où cette phrase a été supprimée, ne transmet pas du tout cet effet de mise en valeur. La brève phrase sans verbe fini ajoute aussi un effet d’oralité au texte source – surtout grâce à l’expression utilisée, ‘Pas une fois.’, qui nous semble très caractéristique du langage oral. Le fait de réduire cette phrase donne à la traduction finnoise un style plus littéraire et l’éloigne ainsi de la perspective scéno-centrique adoptée par l’auteur du texte source. Il est vrai qu’une traduction littérale, ‘*En kertaakaan.*’, ne marcherait pas très bien dans ce contexte. En revanche, il serait possible de traduire cette expression par exemple par ‘*En sitten yhtään.*’ (‘pas du tout’) qui véhiculerait au moins partiellement les effets stylistiques de la phrase source. Comme l’omission de la phrase sans verbe fini n’est pas obligatoire ici, et comme le fait de la supprimer appauvrit le passage considérablement du point de vue stylistique, la stratégie adoptée par le traducteur ne semble pas justifiée.

Après cette séquence, il y a une autre phrase sans verbe fini, ‘Un temps.’ (entre parenthèses, ligne 04). Cette phrase est traduite ‘*tauko*’ (ligne 02), sans changer la construction syntaxique, ce qui semble évidemment être la seule stratégie possible ici. Le mot *tauko* (‘une pause’) est cependant écrit avec une minuscule, et le point apparaissant dans le texte source n’est pas non plus inclus dans la version finnoise. Il y a une occurrence identique dans l’exemple (4) (ligne 04). À première vue, le rôle de ces phrases ainsi que les changements dans leur traduction peuvent sembler triviaux. Néanmoins, à notre avis, les didascalies – qui prennent souvent la forme averbale – sont intéressantes aussi, puisqu’elles sont liées au caractère propre des pièces de théâtre. En effet, ‘Un temps.’ constitue un indice prosodique essentiel et évoque la relation étroite entre les pièces de théâtre et l’oralité. De même, il fait référence au déroulement de l’action sur scène. Les faits d’écrire cette séquence avec une majuscule et de la terminer par un point lui attribuent un statut de phrase autonome ayant une fonction dans le texte. Cela met en valeur aussi son importance situationnelle. Quant à la traduction finnoise, il ne s’agit pas d’une phrase

autonome mais d'un mot isolé écrit entre parenthèses. Cette présentation typographique minimise son rôle dans le texte. D'un autre côté, ce qui est intéressant dans ces occurrences, c'est qu'elles nous permettent de découvrir une cinquième stratégie de traduction des phrases sans verbe fini : le fait de priver la phrase source de son statut de phrase (en supprimant la majuscule et le point). Cette stratégie concerne uniquement les didascalies.

L'extrait ne comporte pas d'autres occurrences de phrases sans verbe fini.

Le dernier exemple présente encore une occurrence d'un cas où une phrase sans verbe apparaissant dans le texte français n'a pas été traduite. L'extrait a été tiré des *Mains sales*.

(8) TEXTE ORIGINAL (EN FRANÇAIS) :

01 Hugo : *Il faut me croire, je t'en supplie.*

02 Jessica : *Je te croirai si tu crois que je suis sérieuse.*

03 Hugo : *Bon. Eh bien, je te crois.*

04 → Jessica : Non. *Tu joues à me croire.*

05 Hugo : *Nous n'en sortirons pas. (On frappe à la porte.) Entrez !*

(Les mains sales. Sartre 1948 : 75–76).

(8) TRADUCTION (EN FINNOIS) :

01 Hugo: *Sinun täytyy uskoa minua.*

02 Jessica: *Uskon, jos sinä uskot, että olen vakavissani.*

03 Hugo: *Hyvä on, minä uskon.*

04 → Jessica: *Nyt sinä leikit uskovasi minua.*

05 Hugo: *Mehän olemme umpikujassa. (Ovelle kolkutetaan.) Sisään.*

(Likaiset kädet. Sartre 1948 : 39. Traduit par Toini Kaukonen en 1966).

Dans cet exemple, la phrase sans verbe fini 'Non.' (ligne 04) n'est pas incluse dans la traduction finnoise. La réplique de Jessica 'Non. Tu joues à me croire.' est traduite '*Nyt sinä leikit uskovasi minua.*' (ligne 04), 'maintenant tu joues à me croire', sans le 'non' apparaissant au début de la réplique dans le texte original. La phrase sans verbe fini ne pourrait pas être conservée ici, puisque la négation du finnois est un verbe. Le fait de la supprimer complètement ne serait cependant pas obligatoire : le 'non' pourrait très bien être traduit directement '*Ei.*' en conservant la fragmentation de la réplique ou bien, il pourrait être relié à la phrase suivante à l'aide d'une virgule. Ici, il est évident que le fait de supprimer le 'non' change le ton de la réplique de Jessica. Dans le texte original, Jessica nie directement ce qu'Hugo vient dire (ligne 03), tandis que dans la

traduction, la négation n'est présente qu'implicitement. De plus, dans le texte français, la négation est donc mise en valeur par la fragmentation de la réplique. En effet, Drillon (1991 : 131) remarque que si une courte phrase sans verbe fini se termine par un point, celui-ci donne un 'caractère affirmatif' à la phrase.

A l'oral, le 'non' constituerait une unité prosodique autonome ; le point qui le termine indique typiquement une chute mélodique suivie d'une pause (Lehtinen 2007). Selon Chafe (1988 : 397), les lecteurs de n'importe quel texte éprouvent une image auditive de différents intonations, accents, pauses, rythmes et qualités de la voix. A l'écrit, la ponctuation constitue le moyen principal pour exprimer ces traits prosodiques couverts (*ibid.*). Dans cet exemple, la prosodie conclusive et le rythme saccadé indiqués par la ponctuation créent un ton affirmatif et déterminé. Comme le 'non' est omis dans la traduction, la réplique entière y est privée aussi de ses effets stylistiques.

L'extrait comporte aussi une autre phrase sans verbe fini : 'Bon.' (ligne 03). Dans ce cas, le mot 'bon' (*hyvä*) est traduit *hyvä on* ('c'est bien'), avec le verbe 'être' (*olla* ; PRÉ, 3<sup>e</sup> P. SG.). De plus, la séquence ainsi formée est reliée à la phrase suivante '*minä uskon*' ('je crois') par une virgule : '*Hyvä on, minä uskon.*' (ligne 03) → 'd'accord, je te crois'. Ici, le fait d'ajouter un verbe est lié au choix d'expression utilisée. Le mot *hyvä* peut être utilisé aussi isolément, mais pas dans ce contexte. En revanche, il serait possible d'employer ici l'expression *no hyvä*, qui ne comporte pas de verbe. De toute façon, ce n'est pas la présence du verbe qui est essentielle ici, mais c'est plutôt le changement du rythme de la réplique. Comme dans le cas du 'non' (ligne 04), ici aussi, la réplique originale commence par une unité prosodique très brève, terminée par une intonation conclusive. Dans la traduction, le fait de relier les deux phrases modifie le rythme de la réplique et atténue ainsi l'effet stylistique créé par la phrase sans verbe fini. En effet, dans le texte original, le point apparaissant après la particule 'bon' n'indique pas seulement la fin d'une unité prosodique mais aussi la fin d'une phase dans l'action du personnage : d'abord, Hugo se détermine et ensuite, il le dit à Jessica. Dans la traduction, l'action se présente d'une manière plus continue, dépourvue de toute hésitation.

### 3. Conclusions

En guise de conclusion, on peut constater que les phrases sans verbe fini constituent un trait caractéristique des pièces de théâtre étudiées. En raison

des effets d'oralité et d'expressivité créés par le phénomène (Larjavaara 2003 ; Riegel, Pellat & Rioul 2004), ces phrases constituent un moyen essentiel de dramatisation du texte. Les phrases sans verbe fini sont fréquentes également dans les versions finnoises de ces ouvrages. Néanmoins, il est notable que dans un tiers des cas, la phrase est traduite en finnois sans conserver la construction syntaxique originale. Dans certains contextes, le changement de la construction est justifié par des raisons grammaticales ou par des raisons liées à l'usage idiomatique de la langue (emploi d'expressions figées, pratiques établies de l'interaction réelle, etc.). Néanmoins, notre corpus comporte aussi de nombreux cas où le changement de la construction sans verbe ne serait pas nécessaire, mais il semble être effectué uniquement dans le but de limiter le nombre total des phrases de ce type. D'un côté, cela suggère que la construction sans verbe fini constitue un trait plus marqué dans le texte finnois que dans le texte français. (Pour confirmer cette hypothèse, il faudrait cependant étudier une grande quantité de textes différents.) De l'autre côté, cela peut s'expliquer par une tendance normative des traducteurs à une écriture texto-centrique. En effet, comme les phrases sans verbe fini servent à la dramatisation de la parole, leur réduction a souvent pour conséquence un changement de la perspective : la perspective scéno-centrique adoptée par Sartre est remplacée dans la traduction finnoise par une perspective texto-centrique. De plus, les traducteurs ont tendance à la standardisation du texte (*cf.* la loi de standardisation de Gideon Toury : Toury 1995, 2004).

Trois stratégies différentes sont employées par les deux traducteurs pour interpréter les phrases sans verbe fini sans conserver la construction originale. L'ordre de fréquences de ces différentes stratégies est le même dans les deux ouvrages traduits, bien que le traducteur ne soit pas le même. La stratégie la plus fréquente consiste à ajouter un verbe fini à la phrase. La phrase sans verbe fini peut également être reliée à la phrase précédente ou à la phrase suivante qui peut être soit une phrase verbale soit une autre phrase sans verbe fini. Il y a aussi un certain nombre d'occurrences où la phrase sans verbe fini apparaissant dans le texte français a été complètement omise par le traducteur.

Le problème qui se pose lorsque la construction sans verbe fini n'est pas conservée est que le changement de la structure atténue souvent les effets stylistiques véhiculés par la phrase source. Grâce à leur brièveté, les phrases sans verbe créent un rythme évoquant le langage oral. Les faits d'ajouter un verbe ou de relier des phrases changent le rythme de la réplique et diluent ainsi l'effet d'oralité. Une autre caractéristique des

phrases sans verbe fini est que leur interprétation est le plus souvent complètement dépendante du contexte. Par conséquent, dans les pièces de théâtre, ces phrases mettent en valeur le fait que les dialogues soient destinés à être joués sur scène où l'ancrage situationnel des répliques est assuré par de nombreux moyens multimodaux. En plus des répliques des personnages, les didascalies qui donnent des indications à la régie et aux acteurs prennent souvent la forme averbale.

Souvent, un verbe est ajouté à la traduction pour expliciter l'action du personnage. D'un côté, l'avantage de ce procédé est qu'il sert à éviter des ambiguïtés. D'un autre côté, l'explicitation diminue la nécessité du lecteur de se mettre à la situation fictive et d'imaginer les événements de la pièce tels qu'ils se dérouleraient sur scène. De ce fait, le changement de la structure peut éloigner les répliques du style qui est propre aux pièces de théâtre. Il en va de même pour les cas où le texte original relève deux phases de l'action consécutives, et ces phases sont unies lorsque les phrases sont reliées.

Dans beaucoup de cas, la construction sans verbe fini constitue un procédé de mise en relief d'un élément de la réplique. Il est évident que l'effet de mise en relief disparaît si la construction sans verbe n'est pas conservée ou si l'élément en question est complètement supprimé. De même, les phrases sans verbe fini sont souvent plus expressives que les phrases verbales. Ainsi, l'effacement de ces phrases a pour conséquence de diminuer le niveau d'expressivité de la réplique concernée. Lorsqu'il s'agit de phrases exclamatives, cet effet de diminution d'expressivité est souvent secondé par un changement de la ponctuation.

## Références

- Catach, Nina (1996) *La ponctuation*. Paris : PUF.
- Chafe, Wallace (1988) Punctuation and the prosody of written language. *Written Communication* 5(4) : 396–426.
- Chesterman, Andrew (1997) *Memes of Translation*. Amsterdam/Philadelphia : John Benjamins.
- Combettes, Bernard & Kuyumcuyan, Annie (2010) Les enjeux interprétatifs de la prédication averbale dans un corpus narratif : énoncés nominaux et représentation fictionnelle de processus énonciatifs et cognitifs. *Discours* 6, <<http://discours.revues.org/index7703.html>> (le 30 octobre 2013).
- Costaouec, Denis (2009) Pour une approche fonctionnaliste. In Françoise Guérin & Denis Costaouec (éds), *Dynamique et changement en syntaxe. Études de cas*, pp. 15–57. Fernelmont : Éditions Modulaires Européennes.
- Drillon, Jacques (1991) *Traité de la ponctuation française*. Paris : Gallimard.

- Gambier, Yves (2008) Stratégies et tactiques en traduction et interprétation. In Gyde Hansen, Andrew Chesterman & Heidrun Gerzymisch-Arbogast (éds), *Efforts and Models in Interpreting and Translation Research*, pp. 63–82. Amsterdam/Philadelphia : John Benjamins.
- Gardes-Tamine, Joëlle (1990) *La Grammaire (2, Syntaxe)*. Paris : Armand Colin.
- Hakulinen, Auli ; Vilkuna, Maria ; Korhonen, Riitta ; Koivisto, Vesa ; Heinonen, Tarja Riitta & Alho, Irja (éds) (2004) *Iso Suomen Kielioppi*. Helsinki : SKS.
- Havu, Eva (2009) Prédications averbales en finnois : subordination ou autonomie ? *Cahiers d'études hongroises* 15 : 207–216.
- (2010) Équivalents finnois des prédications averbales premières et secondes. *Discours* 6, <<http://discours.revues.org/index7718.html>> (le 30 octobre 2013).
- Helasvuo, Marja-Liisa (1997) *When discourse becomes syntax: Noun phrases and clauses as emergent syntactic units in Finnish conversational discourse*. Thèse de doctorat. Santa Barbara : University of California.
- Kalmbach, Jean-Michel & Sundelin, Seppo (2000) [2008] *Suomi–ranska–suomi-sanakirja*. Helsinki : WSOY.
- Larjavaara, Meri (2003) Fixer le regard : Phrases sans verbe fini dans la presse. In Juhani Härmä (éd.), *Le langage des médias : discours éphémères ?* pp. 55–66. Paris : L'Harmattan.
- Lefevre, Florence (1999) *La phrase averbale en français*. Paris : L'Harmattan.
- (2007) Le segment averbal comme unité syntaxique textuelle. In Michel Charolles, Nathalie Fournier, Catherine Fuchs & Florence Lefevre (éds), *Parcours de la phrase. Mélanges en l'honneur de Pierre Le Goffic*, pp. 143–158. Paris : Ophrys.
- Lehtinen, Mari (2007) L'interprétation prosodique des signes de ponctuation – L'exemple de la lecture radiophonique de *L'Étranger* d'Albert Camus. *L'Information grammaticale* 113 : 23–31.
- (2011) La traduction finnoise des phrases sans verbe fini dans *Les Mains sales* de Jean-Paul Sartre. In Bogdanka Pavelin Lešić (éd.), *Francontraste 2 : La francophonie, vecteur du transculturel*, pp. 217–227. Mons : Éditions du CIPA.
- Malone, Joseph L. (1988) *The Science of Linguistics in the Art of Translation: Some Tools from Linguistics for the Analysis and Practice of Translation*. New York : State University of New York Press.
- Molina, Lucia & Hurtado Albir, Amparo (2002) Translation techniques revisited. A dynamic and functional approach. *Meta* 47(4) : 498–512.
- Newmark, Peter (1988) *A Textbook of Translation*. New York : Prentice Hall.
- Noailly, Michèle (2002) L'ajout après un point n'est-il qu'un simple artifice graphique ? In Jacqueline Authier-Revuz & Marie-Christine Lala (éds), *Figures d'ajout : phrase, texte, écriture*, pp. 133–145. Paris : Presses Sorbonne Nouvelle.
- Ono, Tsuyoshi & Thompson, Sandra A. (1994) Unattached NPs in English Conversation. In *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: General Session Dedicated to the Contributions of Charles J. Fillmore*, pp. 402–419, <<http://elanguage.net/journals/bls/article/viewFile/2989/2925>> (le 30 juin 2014).
- Pym, Anthony (2010) *Exploring translation theories*. London / New York : Routledge.

- Rabatel, Alain (2000) Valeurs représentative et énonciative du « présentatif » *c'est* et marquage du point de vue. *Langue Française* 128 : 52–73.
- Riegel, Martin ; Pellat, Jean-Christophe & Rioul, René (2004) *Grammaire méthodique du français*. Paris : PUF.
- Sacks, Harvey ; Schegloff, Emanuel & Jefferson, Gail (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4) : 696–735.
- Sartre, Jean-Paul (1947) *Huis clos*. Paris : Gallimard.
- (1947) [1966] *Suljetut ovet*, traduit par Marja Rankkala. Helsinki : Otava.
- (1948) *Les mains sales*. Paris : Gallimard.
- (1948) [1966] *Likaiset kädet*, traduit par Toini Kaukonen. Helsinki : Otava.
- Traverso, Véronique (1999) *L'analyse des conversations*. Paris : Nathan.
- Toury, Gideon (1995) *Descriptive Translation Studies and beyond*. Amsterdam/ Philadelphia : John Benjamins.
- (2004) Probabilistic explanations in translation studies. Welcome as they are, would they qualify as universals? In Anna Mauranen & Pekka Kujamäki (éds), *Translation Universals. Do they Exist?* pp. 15–32. Amsterdam/ Philadelphia : John Benjamins.
- Tuomarla, Ulla (2004) La parole telle qu'elle s'écrit ou la voix de l'oral à l'écrit en passant par le discours direct. In Juan Manuel López Muñoz, Sophie Marnette & Laurence Rosier (éds), *Le discours rapporté dans tous ses états*, pp. 328–334. Paris : L'Harmattan.
- Vázquez-Ayora, Gerardo (1977) *Introducción a la traductología*. Washington DC : Georgetown University Press.

### Liste d'abréviations utilisées

- ENC : enclitique  
 IMP : imparfait  
 PRÉ : présent  
 P : personne  
 SG : singulier  
 PL : pluriel  
 NÉG : forme négative  
 ZÉRO : personne zéro

Coordonnées de contact :

Mari Wiklund  
Département des langues modernes  
B.P. 24  
00014 Université de Helsinki  
Finlande  
Courriel : mari(dot)wiklund(at)helsinki(dot)fi



Peter Wikström

## **#srynotfunny: Communicative Functions of Hashtags on Twitter**

### **Abstract**

This study investigates various communicative functions served by hashtags in written communication on Twitter from a linguistic pragmatic perspective. A tweet containing a hashtag links to, and is integrated into, a timeline of other tweets containing the same hashtag. Thus, hashtags are by default categorizing or organizing; a user of Twitter may add the tag *#food* to their tweet to integrate it into a general conversation about this topic. However, this study demonstrates that hashtags are also used creatively to perform other communicative functions. In the data presented, hashtags are employed as complexly multifunctional linguistic devices for, among other things, structuring information, playing games, and engaging in reflexive meta-commentary. Notably, while pragmatic methodology is typically applied to speech, this study indicates that a traditional speech acts framework may be profitably applied to written communication in new media.

### **1. Introduction**

The present study investigates various communicative functions served by *hashtags* as a feature of written communication on the social network site (SNS) Twitter from a linguistic pragmatic perspective. SNSs have become an important part of the online social and linguistic domain, and Twitter is one of the dominant players at present. By 2012, the network had around 530 million registered users, posting around 175 million messages per day (Basch 2012; Honigman 2012). As a social network site (cf. boyd & Ellison 2008), Twitter permits users to post public or semi-public *tweets* – individual messages of 140 or fewer characters – and follow the postings of other users.

Twitter has as yet received relatively little attention from linguists, unlike other forms of computer-mediated communication (CMC) such as

email, IM (instant messaging), blogging, and SMS text-messaging, and linguistically relevant work on Twitter to date tends toward approaches that can very roughly be described as sociological and toward a preference for large-scale computational methods. Other studies focus less on what is said and how, instead focusing on who is talking to whom. That is, the focus is often on mapping the topological structure of social networks as in terms of who “follows” or “mentions” whom, how communities of sorts arise around particular popular users, by which channels information is diffused, and so forth (see e.g. Weng, Lim, Jiang & He 2010; Wu, Hofman, Mason & Watts 2011; Murthy 2011; Murthy 2013). Other studies are of relevance to political science, economics, or journalism. For example, Jansen, Zhang, Sobel and Chowdury (2009) describe Twitter as a medium for consumer word of mouth, Kwak, Lee, Park and Moon (2010) analyze Twitter as a news medium and O’Connor, Balasubramanyan, Routledge and Smith (2010) study the correlation between public opinion polling and sentiment expressed contemporaneously on Twitter. Little CMC research on Twitter seems to be ‘purely’ linguistic in nature, in the sense of focusing mostly or exclusively on the linguistic structure of Twitter discourse. Studies that take linguistic structure into account are often sociolinguistic, using “linguistics as a lens on community” (Zappavigna 2012: 10; cf. also Seargeant & Tagg 2014). The present study contributes to linguistic work on Twitter by focusing on close qualitative analysis of language use at the utterance level.

More specifically, the present study focuses on Twitter’s hashtag feature. Hashtags are hyperlinks generated by the user prefixing a string of letters in a tweet with a hash symbol (#), for example *#food* or *#thatshowyoudoit*. A tweet containing a hashtag links to, and is integrated into, a timeline of other tweets containing the same hashtag. Thus, hashtags are by default categorizing or organizing: if a tweet treats the subject of food in some manner, it may include the hashtag *#food*, integrating it into an on-going communal conversation about the topic of food on Twitter.

However, many hashtags appear to have little to do with categorization, but rather seem to be used to perform a variety of other communicative functions. For instance, Huang, Thornton and Efthimiadis (2010) conclude that tagging is “conversational” as opposed to “organizational,” based on large-scale data analysis.<sup>1</sup> The authors are also

---

<sup>1</sup> Page (2012) argues that the “conversational” quality of hashtags may be quite limited e.g. in the discourse of celebrities or corporations on Twitter. See also Honeycutt &

concerned with the meaning of the tags in context, but because their qualitative analyses are limited only to the most common tags in their dataset, they tend not to find less commonly used tags with more situationally dependent pragmatic functions of the kind that the present study is primarily concerned with. They do, however, identify what the present study calls “hashtag games,” labeling them “micro-memes” (Huang et al. 2010: 3). Zappavigna (2011) uses a Systemic-Functional Linguistic approach to analyze community building by means of evaluative language in tweets. She finds that topic marking hashtags are often used to indicate the target of appraisal, as, for instance when tweets evaluating President Barack Obama contain the hashtag *#Obama* (Zappavigna 2011: 12). Zappavigna (2012, 2014) shows how hashtags serve to facilitate what she terms “searchable talk” and “ambient affiliation,” by permitting users to align their experiences or attitudes with those of other users on the same or similar topics, and how hashtags maybe used to mark various kinds of interpersonal bonds.<sup>2</sup>

### 1.1 Aims and data

The primary aim of the present study is to identify and analyze a broad range of communicative functions served by hashtags on Twitter using a linguistic pragmatic framework based on speech act theory. A secondary aim is to assess whether this analytical framework, typically employed for the analysis of spoken interaction, may profitably be applied to typewritten language data in a CMC environment. This secondary aim is motivated by a general scholarly discussion within the field of CMC about the supposed or actual spoken-likeness of various online platforms for typewritten communication (cf. e.g. Crystal 2006; Baron 2008, 2009; Dresner & Herring 2010; Jonsson 2013). The analytical framework is presented in section 2. For the analysis (section 3), 72 illustrative examples were selected from an initial sample of approximately 1200 tweets containing hashtags. The dataset was compiled in November 2010, via manual collection from the timelines of arbitrarily selected users and, in a few

---

Herring (2009), on uses of the @-addressivity device on Twitter, which may more readily afford conversation than do hashtags.

<sup>2</sup> There is also a meta-functional analysis of hashtags forthcoming from Zappavigna (personal communication), which provides further insight into the communicative complexity of hashtags from a Systemic-Functional perspective.

cases, via manual searches for specific hashtags.<sup>3</sup> The sole criterion for inclusion was the presence within the body of the tweet of a hashtag that was not plainly a simple topic marker (with the exception of a few examples meant to illustrate the topic marker function). Neither the sampling procedure nor the selection of examples presented below were intended to yield a quantitatively representative picture of hashtag usage on Twitter as a whole, but rather to identify and exemplify a wide variety of functions that tagging can be used to perform.

## 2. Analytical framework

The communicative functions performed by the hashtags in the dataset are treated in separate subsections throughout section 3, according to the following categories: 3.1 *topic tags*, 3.2 *hashtag games*, 3.3 *meta-comments*, 3.4 *parenthetical explanations/additions*, 3.5 *emotive usage*, 3.6 *emphatic usage*, 3.7 *humorous and playful usage*, and 3.8 *memes and popular culture references*. These categories were arrived at during the collection of the dataset, and then refined upon closer analysis. The categories are ad hoc and should primarily be considered as heuristic. That is, the sorting of examples of hashtag usage into distinct categories of communicative functions is primarily meant to organize the presentation of the findings and to emphasize the variety of uses to which hashtags are put, rather than to constitute a taxonomy in any proper sense. Further, the categories should not be understood as mutually exclusive. Multifunctionality may be the norm rather than the exception.

In the analyses presented in section 3, the communicative functions are explicated and discussed in terms of what may be called a “traditional” pragmatic speech act theory framework based in the work of John Austin, John Searle, and Paul Grice. The concluding discussion in section 4 presents, *inter alia*, a summary of how the communicative functions may be characterized in terms of the pragmatic concepts introduced here.

Below, the tweets presented are analyzed as speech acts or illocutionary acts intended by their authors to express one or multiple kinds

---

<sup>3</sup> Twitter’s terms of service make it clear to its users that public tweets and profile information will be openly available (Twitter 2014). Accordingly, no data were collected that were not voluntarily made public. However, the user names of anyone other than celebrities, public figures, and institutions have been anonymized as @user. Beyond anonymization, the contents of the presented tweets have not been altered in any way.

of specific *illocutionary force* (Searle 1969; Austin 1976). A more generalized and abstract construal of such force is what Searle terms *illocutionary point*. The illocutionary force of the speech act “What is your name?” is that the speaker wishes to learn from the hearer what the hearer’s name is, and its illocutionary point is a request for information. An utterance can have several illocutionary points, as we often do several things at once in the same utterance (Searle 1979: 29).

Grice argues that we generally infer an interlocutor’s intention by implicitly assuming the observance of a cooperative principle, which he states as follows (Grice 1975: 45):

Make your conversational contribution such as it is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

The principle is an unspoken rule which seems to be taken for granted by the participants in most conversational interactions. Grice’s cooperative maxims, which further explicate what constitutes cooperation, can be summarized as follows (Grice 1975: 45–46):

1. *Quantity*: The speaker is expected to be sufficiently informative, saying neither too much nor too little to be properly understood.
2. *Quality*: The speaker is expected to be truthful to the best of his/her ability.
3. *Relation*: The speaker is expected to keep his contribution relevant to the present interaction
4. *Manner*: The speaker is expected to strive for lucidity and orderliness in expression.

Grice discusses several ways in which speakers might fail to observe a maxim (Grice 1975: 49), one of which is *flouting*. Flouting occurs when the speaker blatantly and purposefully fails to observe a maxim, thereby exploiting it to generate additional meaning. Throughout interactions, hearers use contextual cues or background knowledge to facilitate the process of inferring the point and force of utterances.

It should be noted that the framework employed in this study is what may be termed traditional pragmatics or speech act theory, and that the field has seen substantial developments since the time of Grice, Austin, and Searle. The employment of a traditional framework should not to be considered a dismissal of later perspectives on speech act theory and cooperation (e.g. Leech 1983), “neo-Gricean” ventures (e.g. Levinson

1983), cognitive approaches to relevance (Sperber & Wilson 1995), or, indeed, the application of pragmatics to CMC (e.g. Yus 2011).

### 3. Analysis and results

The following subsections address the communicative functions outlined above. Each section comprises several sets of examples chosen partly to illustrate the identified functions, but partly, also, to illustrate overlaps between functions and other complexities. As mentioned, the organizing or categorizing function seems to be the default or ‘expected’ mode of usage given how hashtags work technically on Twitter; that is, as hyperlinks to a timeline of tweets containing the same tag. Therefore, the functions Topic tags (3.1) and Hashtag games (3.2), which are most in line with this expected mode of use, are presented first, so that they may serve as a form of baseline for the presentation of the arguably more functionally complex uses in the following subsections.

#### 3.1 Topic tags

“Topic tags” is the label suggested in this study for the more or less straight-forward use of hashtags to integrate a tweet into a conversation on a given topic, as in examples (1–4).

- (1) #Golf Tiger Woods’ new home has practice facility. <http://bit.ly/i4ciY2>
- (2) @user hmmn. Yes, the right amount of quaint & old-fashioned. If one has lemon juice & sugar, of course. #pancakes
- (3) Charles #Darwin published Origin of Species on this day in 1859. <http://bit.ly/i8BilW> #evolution #science
- (4) Today’s #TED talk, from #TEDxDU: Mind your matter! Kim Gorgens on brain injury <http://on.ted.com/8ia6>

Examples (1) and (2) are relatively simple: the tweet treats a topic designated by the hashtag. Example (3) is from a user account that posts links to articles and blog posts on the *Encyclopædia Britannica* website, and uses tags relevant to the topic of the linked webpage. In (4), the hashtags #TED and #TEDxDU tie the tweet to the topic of the TED (Technology Entertainment and Design) set of conferences and the University of Denver sub-conference. Examples (5) and (6) function similarly, but are slightly more complex.

(5) This is the worst thing I have ever heard. #amas #blackeyedpeas

(6) Good game, #Chiefs

In (5), the hashtag usage can be unpacked pragmatically as follows: 1. The poster makes deictic reference to something being “the worst thing I have ever heard.” 2. The reader presumes that the cooperative principle is being adhered to. 3. For the maxim of quantity to be observed, there must be some additional information clarifying what is being referred to deictically, and for the maxim of relation to be observed, the tacked on hashtags at the end must be relevant to the utterance. 4. The reader thus supposes that “this” refers to “#amas” and “#blackeyedpeas.” 5. Given the requisite background knowledge that “the AMAs” is a common way to refer to the American Music Awards and that there is a popular band called the Black Eyed Peas, the reader concludes that the poster is commenting on the performance of the band at the awards show. The tags supply the reader with information regarding what is being referred to in the main text of the tweet, which makes them similar to many of the hashtags discussed in section 3.4 on parenthetical additions or explanations. However, it seems that the primary communicative function of the hashtags in this tweet is to integrate the tweet into an ongoing public conversation about the AMAs and the Black Eyed Peas (cf. Zappavigna 2012). Example (6) is felicitously interpreted along the same lines, with the tag making the tweet part of a conversation about the sports team called Chiefs.

Example (7) below differs from the previous examples in that it was intended to coordinate multiple tweets by one user rather than to coordinate multi-user interaction.

(7) I’m going to post, now, a dozen or so Tweets that tell a happy story that really happened. #happystory

Example (7) was indeed followed by “a dozen or so” tweets that relayed the “happy story” in brief chunks. These chunks were all tagged #happystory, in order to mark that they belonged to this topic. Examples (8–11) below are posts taken from the timeline of the hashtag #MayTheForceBeWithKatie, which derives from the widely recognized *Star Wars* incantation “May the force be with you!” The posts refer to a news story about a young girl who was bullied for having a *Star Wars* themed water bottle.

(8) Unacceptable! RT: First grader bullied for having a Star Wars water bottle. <http://tinyurl.com/22w6fzz#MayTheForceBeWithKatie>

- (9) Don't let those boys get you down [#maytheforcebewithkatie](#)
- (10) This is sad...kids can be so mean  
<http://bit.ly/cNT2Pa#starwars#MayTheForceBeWithKatie>
- (11) [#MayTheForceBeWithKatie](#)

Examples (8–10) contain links, summaries, and personal reflections pertinent to the story, while (11) comprises merely the topic tag *#MayTheForceBeWithKatie*. Examples (8–10) are easily understood as cooperative contributions to an ongoing conversation, but (11) fails to observe several maxims. It is terse (quantity), and obscure (manner), and consequently of questionable relevance to the conversation. However, a reader assuming the cooperative principle can infer that the repetition of the hashtag without comment is intended to function something like the chanting of a slogan. It has the illocutionary force of expressing support for a related idea or cause, and also brings the cause to the attention of others by making noise about it. Accordingly, it appears that *#MayTheForceBeWithKatie* performs an implicit emotive or emphatic function in addition to its topic tagging function.

### 3.2 Hashtag games

The use of hashtags which is here called hashtag games is not markedly different from topic tagging, as the basic function performed by the tags is still organizing or categorizing. Here, however, the purpose of the tag is not to mark a tweet as belonging to a certain conversation or relating to a certain topic, but rather to mark it as participating in an ongoing communal game (cf. Huang et al. 2010), often grounded in some kind of word play. One such game is exemplified by (12–14), which are contributions to the timeline of the hashtag *#BoringPrequels*.

- (12) [#boringprequels](#) The Empire Holds a Planning Meeting
- (13) Earth Trek [#BoringPrequels](#)
- (14) Some Like it Tepid [#BoringPrequels](#)

This is a common type of game played on Twitter, with a simple premise contained within a unique hashtag. The object of the game is to make up the title of a humorously boring prequel to an existing film. Examples (12–14) are all by different posters, but follow the same formula, whereby the imagined prequel title is closely reminiscent of the original title while

adding a new twist to it. Example (12) is based on *Star Wars episode V: The Empire Strikes Back*, (13) on *Star Trek*, and (14) on *Some Like It Hot*. Below, examples (15–17) show a variant of the same game-type.

- (15) Indifference Island [#DullRealityShows](#)
- (16) World's Quietest Libraries [#DullRealityShows](#)
- (17) America's Next Top Insurance Adjuster [#DullRealityShows](#)

As different games of this type tend to be formally and conceptually similar to each other, any experienced user of Twitter encountering such a tweet should be able to pick up on it and contribute if she so desires.

Examples such as (18–20) are more like conversational contributions to a specifically themed topic than like rule-bound wordplay, but do share the formulaic quality of the previously exemplified games.

- (18) [#rememberwhen](#) gas was a 1.50..a bottle of coke was 1.00 and a bag of chips was 25 cents
- (19) [#rememberwhen](#) Mtv was cool and just played music videos!!!!
- (20) [#RememberWhen](#) vampires used to explode in the sun, not sparkle like fairies?

The tag [#RememberWhen](#) serves the same purpose of establishing a simple but recognizable and unique premise, and being part of a template that must be adhered to for successful participation. The point of [#RememberWhen](#) is for the poster to “fill in the blank” with a nostalgic or mock-nostalgic memory. A reader who is assuming cooperation and is familiar with this conventional form of indirect speech act will infer that the illocutionary point of the utterances is not to request information (e.g. “Do you actually remember when MTV played mostly music videos?”) but rather to make a statement or express an opinion, specifically with the force of expressing preference for the past situation. Below, examples (21–23) illustrate a game similar to [#RememberWhen](#).

- (21) [#lemmeguess](#) You not who I think you are right?
- (22) [#lemmeguess](#) you love me but you can't be with me.?
- (23) [#lemmeguess](#) thats ah real polo shirt??then why instead of the man on the horse you got ah nigga walkin ah dog

This game appeared to primarily attract African American participants, judging by profile pictures and African American Vernacular English

(AAVE) linguistic features throughout the timeline. Indeed, the spelling of the organizing hashtag of the game might be seen as signaling that this game in some sense belongs to an AAVE community on Twitter. Accordingly, it seems that this tag performs an additional social function beyond organizing the game.

### 3.3 Meta-comments

Even though there is a slight difference between hashtags that relate a tweet to a topic of conversation and hashtags that mark a tweet as a contribution to a game centered on that tag, both types of usage have in common that the technical functioning of hashtags on Twitter (as hyperlinks) is relevant or even crucial to them. However, a lot of hashtag usage seems not to be as concerned with that functionality, as in the following:

(24) *.@user* You don't look a day more than 12 parsecs! *#YesIKnow*

For a reader lacking the requisite background knowledge, it might seem that the tag *#YesIKnow* is not relevant to the rest of the tweet, meaning that the poster is failing to observe the maxim of relation. However, a reader who recognizes the common mistake of using “parsecs” as if it were a measurement of time (when it is actually an astronomical unit of length) and notices that the poster is someone who is likely to be well aware of this common mistake (in this case, the astronomer and blogger Phil Plait, *@BadAstronomer*) will likely interpret *#YesIKnow* as a tongue-in-cheek acknowledgement of the error. Thus, the illocutionary point of the hashtag is to make a meta-comment about the main content of the tweet, with the specific force that the apparent terminological mistake in the tweet is intended ironically.

The tags in examples (25) and (26) below are also meta-comments, but ones which seem to have the purpose of explicitly stating the illocutionary point of the tweet:

(25) Also think “Webcomics” should only be used to refer to a spec. biz model. It's an outdated descriptor for a genre or a community. *#opinion*

(26) *#statement*: I'm cooler than you

Example (25) is clearly expressing an opinion, so it seems reasonable to conclude that the tag is meant as an acknowledgement of this. Of course, since it is clear from the main text of the tweet that an opinion is being

expressed (signaled by the opening, “Also [I] think [that]...”), the tag is superfluous, in breach of the maxim of quantity. The tag may be interpreted as flouting the maxim of quantity in order to generate pragmatic force as a hedge. The tag functions as a disclaiming meta-comment suggesting that the utterance is not intended as a definitive statement of fact. Example (26) similarly comprises a declarative clause and a tag that superfluously describes what kind of utterance it is, but emphasizing that a statement as a statement is not likely to have the disclaiming effect that labeling an opinion as an opinion does. Rather, the intended effect seems to be the opposite: to enhance the force of the utterance as an affirmation of undeniable fact. However, given the triviality and subjectivity of the topic (someone’s relative “coolness”), and the lack of a specific addressee, the force of the tag as an affirmation of factuality is presumably intended ironically. The poster would thus be flouting the maxim of quality by claiming more than the poster has warrant for in order to perform a joke.

In (27–29), the tags *#itsajoke* and *#justkidding* are explicitly clarifying utterance intention:

- (27) Has Rooney shagged Katie weasels nan ??#itsajoke
- (28) @user1 likes @user2 's cock :b LMFAO #justkidding #itsajoke #dntgetmad #imsrry #lmao
- (29) @user LOL WAY TO GO BILLIAM NOW I HAVE TO HATE U 4EVR. #justkidding #idontactuallytalkinallcaps

The tag *#justkidding* in particular is reminiscent of common online linguistic conventions of disambiguation, such as putting words or phrases within asterisks (e.g. \*smiles\*), or adding emoticons or abbreviations. In (28), this connection between hashtag usage on Twitter and general online paralanguage is further established by the tag *#lmao*, which is the abbreviation of the phrase “laughing my ass off” turned into a hashtag.

The tag *#idontactuallytalkinallcaps* (“I do not actually talk in all capitals”) in (29) is also interesting. Writing in capitals is often perceived online as a way of being loud and obnoxious. Consequently, the poster evidently feels compelled to clarify that she is not “actually” behaving in this manner, but is merely staging the behavior. Thus, a meta-comment tag can perform the function of maintaining one’s public self-image (*face* in pragmatic terminology) by directly cancelling undesired pragmatic force. Example (30) further illustrates this function:

(30) On 9/11 the world united in horror&despair. Let's not wait for tragedy to be united. #srynotfunny

Few readers would presumably expect a tweet about the event of 9/11, posted on the anniversary of the event, to be an attempt at humor. The implicit acknowledgment of such an expectation does make sense, however, given that the poster is the comedian Sarah Silverman (@SarahKSilverman). As a comedian, Silverman is writing for an audience with certain expectations. Thus, again, the meta-comment hashtag is utilized as a device for dealing with public self-image, clarifying utterance intention by the cancellation of undesired force.

### 3.4 Parenthetical explanations and additions

Examples (31–33) below have in common that the hashtags constitute additions that explain the utterance meaning of the main text of the tweets.

(31) I am being held hostage by this Q.#stillplayingscrabble

(32) This is why I've been offline today. <http://yfrog.com/n6v9ssj>  
#ILoveHikingInColorado #ButNo3GInMountains

(33) Did you start November the 6th like this (holds up bandaged hand to camera and makes a sad face)?#oldfireworksafetyadvert

These tags do not disambiguate utterance intention, like the meta-comment tags described above do, but rather supply additional information to help readers lacking relevant background knowledge make sense of the tweet. Example (31) is about the poster's predicament with the difficult letter Q in an ongoing game of Scrabble. Without the tag, this would not be easy to infer for a reader who had missed any previous mention of the game, as the utterance "I am being held hostage by this Q" alone would constitute a failure to observe the maxim of quantity. In (32), the [yfrog.com](http://yfrog.com/n6v9ssj) link leads to a mobile phone photograph of mountains, and the two hashtags serve to facilitate the inference that the poster was offline while hiking in the mountains. Example (33) is meant to evoke an old public safety advertisement, and the tag is likely intended as a parenthetical explanation to clarify the meaning of the tweet to readers who do not immediately recognize the cliché. These tags thus serve as cooperative additions with the illocutionary point of providing information which helps satisfy the maxim of quantity. The tags are "parenthetical" in the sense that the information they provide is likely superfluous to any reader who already

knows the context (e.g. a reader who has followed the previous tweets about the scrabble game in (31)).

In (34), however, it is more difficult to argue that the hashtag is parenthetical:

(34) My doctor was shocked when I guessed he'd prescribe me Biaxin.#Biaxinpensbiaxincalendarbiaxinnotepad

The force carried by the hashtag, namely the suggestion that the doctor prescribes a certain medicine in exchange for what one might call sponsorship, appears to be the central component of the tweet. Without it, readers would probably feel that the utterance fails to observe the maxim of quantity, as a mere introduction to an anecdote that has no point or resolution. However, with the tag, the utterance becomes a wry observation about the pharmaceutical industry. A reply to (34) displays effectively the functional plasticity of hashtags:

(35) @SarahKSilverman#Biaxinpensbiaxincalendarbiaxinnotepad And they say it's the R&D (not the marketing) that inflates the cost of medications.

This tweet uses the same tag to indicate that it continues on the topic established in (34) (which was posted by @SarahKSilverman). Thus, a hashtag which could not reasonably be construed as a topic tag in its original tweet becomes one when another poster uses it as such.

Examples (36–38) contain parenthetical additions that give non-essential extra information:

(36) I saw two of these dogs today. They are so cute. <http://j.mp/3ypBrb>. #shibainu

(37) Check that sky: <http://twitpic.com/38b7gl> #kansas

(38) It's quite mild tonight. #walkingthedog

These hashtags are not central components of the force of their respective tweets, nor is the information conveyed in them necessary for explaining the text that precedes them. For instance, *#walkingthedog* in (38) is not necessary for enabling any reader to understand the utterance as a casual observation about the weather at the poster's location. Examples (36) and (37) contain links to images of a dog and a landscape, respectively. The hashtags give the breed of the dog and the location of the landscape, as additional information for interested readers.

Examples (39) and (40) show hashtags that add explanation or elaboration of the posters' attitude or feelings.

(39) @user who yu tellin I got the price on the mnu say 4.99 then y I'm spending 7 dollars #Shitainright

(40) Fo Realz, it is blizzarding. #AndiLikeIt

In (39), the poster is expressing dissatisfaction with paying \$7 instead of \$4.99, with the hashtag making explicit and strengthening this force. Example (40) comments on the weather, and the hashtag adds, perhaps as a parenthetical after-thought, that the poster likes it. These tags that make attitudinal additions lead into the emotive and emphatic functions of hashtags treated in the following two sections. This seems especially true of a hashtag such as #*shitainright*, as it is possible that the choice to turn the phrase into a tag was primarily intended to emphasize it as an emotional expression.

### 3.5 Emotive usage

Example (41) below is from the same poster as (39) (#*shitainright*). The tag here performs a similar function, in that it has clear emotionally expressive force:

(41) Dont feel like walking...but ill make it #sigh

However, unlike the tag in (39), it seems that this tag is meant to represent a face-to-face paralinguistic cue. In speech, an utterance like that in (41) might be accompanied by a sigh, strengthening the force of the utterance as an expression of weariness at the thought of having to undertake a hard or tiresome task. Example (42) below features a similar tag:

(42) "We could split the cinnabon!" #HowIMetYourMother #laughs

Here, #*HowIMetYourMother* is a topic tag and a clarification that the quote is from the television show of that name, while #*laughs* is most reasonably construed as a representation of the poster's reaction to what is being quoted. Accordingly, the hashtags here are used to perform expressive functions that might otherwise be performed by writing e.g. <@*user* sighs> or \*laugh\* (cf. Crystal 2006: 38). Many Twitter posters do use the common online typographic markers for this purpose, but apparently some

posters use tagging to the same end. It should be noted that *#laughs* or *#sigh* are not exactly like the paralinguistic cues that they seem to be representing: laughter is often an involuntary reaction, whereas typing *#laughs* requires deliberate intention (cf. Crystal 2006: 37, fn.15). This difference is even more notable when it comes to example (43):

- (43) @user RESPOND TO MY TWEET GODDAMN YOU! ARE YOU DEAD?!  
#Worry #Fret #ShutUpUser

Non-verbal or paralinguistic expressions of worrying and fretting are unlikely to be intentional communicative acts. Accordingly, a tag such as *#Worry*, in addition to being an expression of the poster's emotional state, also carries the illocutionary force that the poster is self-conscious about being worried. The final tag makes this self-consciousness explicit, as the poster is telling himself to "shut up", perhaps reprimanding himself for failing to observe the maxim of quantity by being repetitious, or for failing to observe the maxim of manner by being too loud.<sup>4</sup>

Example (44) also features a tag expressing emotional state, but in a somewhat different manner formally:

- (44) @DesertBus I just sat down at my computer and instinctively typed in the DesertBus url... only to remember it's over. #sadwharrgarbl

Online, "wharrgarbl" originated as the caption of a widely circulated image of a dog trying to drink water from a lawn sprinkler, and is often used to represent ranting or raging incoherence or babbling.<sup>5</sup> In conjunction with the word "sad," as in (44), it is presumably meant to represent some type of gargling throat sound of frustrated dismay. The illocutionary force of the tag seems to be partly emotionally expressive and partly joking.

Other hashtags represent emotionally loaded vocal sounds in a more conventional way:

- (45) Thanks for the retweet @user! For some reason Twitter dropped you off my follow list. That's been fixed! #Grrr
- (46) I woke up to a left leg and knee in horrible pain. I torqued something playing ping-pong yesterday. #ouch

---

<sup>4</sup> The tag contained the first name of the poster, which has here been anonymized.

<sup>5</sup> See <http://www.urbandictionary.com/define.php?term=WHARRGARBL> (accessed 26 September, 2014).

“Grrr” and “ouch” are conventional ways of rendering in writing vocal sounds associated with emotional states. There is also a playful element to the usage. For instance, a growling “grrr” as an expression of anger has something of a cartoonish quality to it, which is enhanced by putting it in a hashtag; a genuinely angry person would presumably not be inclined to engage in such hashtag onomatopoeia.

Examples (47) and (48) below feature yet another way of expressing emotion by means of representation of sound (in this case, manner of pronunciation):

- (47) Omg I can't come down from this praise and worship high from this morning. God's Word is such a healing balm! My spirit needed it! #Jeeesus
- (48) das coold ! i wont get to listen to @user 's cover song til waaaay later . #maaaan /:

If the intention of (47) was to integrate the tweet into a general conversation on the topic or theme of Jesus, it is likely that the tag would have been spelled *#Jesus*. As it stands, a reader assuming that the cooperative principle is being observed, i.e. assuming that the blatant misspelling is purposeful, has to interpret *#Jeeesus* in some other way. Considering the theme of religious ecstasy in the tweet, the tag is most reasonably construed as intended to express an emotional state of fervor. The orthography is likely meant to represent the way in which a vowel might be lengthened in the pronunciation of the name when it is called out in an excited state. Similarly with (48): there is a common way of lengthening the vowel in and giving a falling intonation to the word “man” when it is used in speech to express disappointment, and presumably *#maaaan* is intended to represent this manner of pronunciation. Examples such as these show that it is sometimes difficult to distinguish between the emotive communicative function and the emphatic function treated next.

### 3.6 Emphatic usage

This section deals with tweets using hashtags as an alternative to other forms of emphasis in writing, sometimes in a manner that appears to represent (or at least to be intended to function similarly to) vocal emphasis in speech. Examples (49–52) illustrate this.

- (49) I Bring it on, On my #Own!

- (50) ugh Avis won't rent me a real car because I'm 24 #seriously
- (51) @user Elric only has ONE sword. #Check #Mate.
- (52) @user Lol studyin on a Friday night?It's the weekend #BEFREE lol. & I'm doin nothin:( everyone has work so I had to come home early

In (49), it is difficult to think of reasons compatible with the cooperative principle why the poster chose to turn “own” into a tag. It seems most plausible that the purpose is simply to use the hashtag to set the word apart from the rest of the sentence typographically in order to emphasize it. It is common in writing to mark emphasis by typographically marking letters, for instance by bolding, italicizing, or using upper-case. On Twitter, only the last of these emphatic devices is available. Thus, it is conceivable that many users seize upon the typographic marking created by hashtagging a word or phrase as an alternative. If (49) were spoken, it is easy to imagine how the final word might be given vocal emphasis to strengthen the force of the expression of independence. Similarly, it seems likely that the word “seriously” in a spoken version of (50) would be stressed to express disappointment or incredulity. With (51), it is possible that “check” and “mate” were split up into two separate tags to emulate the way in which one might use stress and rhythm in speech to emphasize the finality of the winning chess move: “*Check...mate.*” In example (52), the poster has also emphasized #BEFREE by putting it in upper-case, presumably to strengthen the force of the imperative.

Example (53) shows particularly expressive use of hashtags for emphasis:

- (53) #why oh why is this guy #kissing a #fucking #RACOON !!!! #nooooooooooooo

Taking into account elements of (53) such as the distraught “why oh why” and the intensifier “fucking,” it seems clear that the main intent with the tweet is to express intense emotion. The tagging is accordingly most reasonably construed as emphatic, thus enhancing the expressive force of the utterance. Of course, there is also a note of playfulness in the tweet, as the horror expressed is so overstated, from the theatrically melodramatic quality of “why oh why” to the excess of hashtags. This flouts the maxims of quantity and quality – overstating the horror such that it cannot reasonably be taken as genuine – presumably for comic effect.

Example (54) shows how emphatic tagging may also be used for distancing, arguably in a manner similar to how scare quotes are used in standard writing.

(54) I hate when hoes be callin me #sir or #mister... stop dat shit  
The force of (54) is that the poster does not consider “sir” and “mister” to be appropriate terms of respect, at least when applied to him by people he considers “hoes”. Thus, it seems that the poster is marking the words by tagging them in a manner similar to how one might in speech utter the words with a dismissive tone of voice (or, again, mark them with scare quotes in writing).

### 3.7 Humorous and playful usage

In what follows, there are several examples of hashtags being employed to fulfill some humorous or playful function, including some examples where hashtags and the practice of tagging itself becomes the subject of playful commentary.

(55) Once again everyone. I do not dye my hair. Not on my head anyway.  
#scarydayglopubes.

Excluding the hashtag, this tweet follows a common joke structure, with a set-up and a punch line. If the cooperative principle were not in force, “Not on my head anyway” would constitute a failure to observe the maxims quantity, relation, and manner. In fact, the poster is flouting these conversational maxims in order to imply, humorously, that he does dye other hair on his body. The hashtag *#scarydayglopubes* adds an additional level of humorous absurdity by suggesting that he dyes his pubic hair with “Day-Glo” color (i.e. pigment that is fluorescent in daylight). The hashtag is thus used to expand the joke with a second punch line.

Similarly, in (56), the illocutionary force of the utterance is generated by flouting maxims.

(56) Having an Alan Partridge moment. Had a protein bar after my shower and got chocolate on my towel. What’ll housekeeping think? #dirtyprotest.

It is suggested that housekeeping staff might mistake some chocolate on a towel for excrement. The tag *#dirtyprotest* furthers the joke, for readers possessing the requisite background knowledge, with a reference to the

1978 “dirty protest” by paramilitary prisoners held in Northern Irish prison facilities, who protested their treatment by smearing excrement on the walls. The tag comically juxtaposes the bit of chocolate on a towel with the conditions of the dirty protest, perhaps suggesting that housekeeping might construe the chocolate on the towel as an act of dirty protest by the poster.

In (57), humor is similarly accomplished by a hashtag that prompts an unexpected inference:

(57) I typed DONG and I meant it. #ownyourtypos

Here, the poster is acknowledging a typo in a previous tweet (*dong* for *don't*), but is making a joke out of it by insisting that it was really what he or she meant. The hashtag *#ownyourtypos* is making a meta-joke: the tag describes what the poster does in the main text of the tweet – he is “owning” the typo (i.e. running with it owning up to it rather than disclaiming it) – but does so in the form of an imperative that sounds like it could be the slogan or catchphrase of a movement. The addition of the hashtag thus prompts the bizarre inference that owning one’s typos is something that one should strive for. Conceivably, the typo “meant” is intentional and meant to add to humorous effect of the tweet.

In (58) and (59), humor arises from how strings of hashtags have been combined:

(58) Because, @user, I don't want them to think I've taken them for granite. #OhYeahIWentThere #NoQuartzWillConvictMe #Rock

(59) Accidentally got the wrong kind of Trader Joe's frozen pizza. #Emergency #Horrors #AbandonHope #BookOfRevelation #NotEnoughCheese

In example (58), the tag *#OhYeahIWentThere* presumably refers to the immediately preceding pun (which plays on the approximate homophony of *granite* and *granted*), with the illocutionary force that the poster is unapologetic for his cringe-worthy geology pun. The second tag, *#NoQuartzWillConvictMe*, continues in this vein, with the poster proclaiming that there will be no legal come-uppance for his punning (even while adding another geology pun to his list of offences). The third tag, *#Rock*, reiterates the geology theme, but looks more like a common topic tag. However, the terseness and sheer matter-of-factness of the hashtag becomes something of a joke itself, in the context of the two preceding tags. The humorous force of the tag *#Rock* is perhaps strengthened by how it flouts the maxims of quantity and relation: Otherwise, since the tweet is

not a productive contribution to the topic of “rock,” the tag is superfluous and irrelevant. Example (59) also features a humorous string of hashtags. The main text of the tweet is a trivial anecdote about the poster getting the wrong kind of frozen pizza. Given the triviality of the problem, the force of the first tag is clearly hyperbolic, and the subsequent tags ramp up the hyperbole with absurd references to Dante’s *Inferno* and the biblical Apocalypse of John. The poster is thus flouting maxims, dramatically overstating his concern, for comic effect. The final tag emphasizes the discrepancy between the tags and the main text of the tweet by apparently noting an especial concern that the frozen pizza he bought might not have enough cheese.

The final examples in this section are comments on excessive hashtag usage, themselves featuring self-consciously excessive hashtag usage:

- (60) Ever perplexed by the particularly Welsh business fixation of #bizarrely #excessive and #obscure #twitter #hashtagging #Flibble
- (61) #HowToGetBlocked #Put #Hash #Marks #On #Random #Words. #Booty #Legs #Hamburgers #Chicken #Sofas #Pillows #ESPN #Grease #Twizzlers
- (62) #i #want #my #own #hash #tag

These examples express stances taken regarding excessive hashtag usage, while themselves demonstrating such excess. Presumably, the self-conscious irony of engaging hyperbolically in the very linguistic behavior that one is condemning is intended to be humorous. Example (61) is perhaps especially noteworthy, since it even uses a mock topic tag, *#HowToGetBlocked*, to summarize its message: “this is how to behave if you want people to block your tweets.” Finally, in example (62), the poster might be using excessive hashtagging to strengthen the force of his expression of desire for a personal hashtag (whatever that might mean), but presumably the poster is simply fooling around with hashtags in a way that is not necessarily meant to signify anything but playfulness (cf. the notion of performing a “ludic self” on social media in Deumert 2014).

### 3.8 Memes and popular culture references

This section deals with a convention that was found among some posters of using Internet memes or other popular culture references (quotes or paraphrases) as hashtags in their tweets. After being coined by evolutionary biologist Richard Dawkins as a cultural analogue for the biological concept

*gene* (Dawkins 2006: 189–201), the term *meme* has come to be used online to refer to phenomena such particular phrases or genres of images or videos that start replicating rapidly throughout an online population.

As may be evident, this notion of memes maps well onto the hashtag games described above. Such games may be considered transient hashtag memes that trend as the game is going on and then fall out of use (cf. the “micro-memes” of Huang et al. 2010). The memes considered in this section, however, are less transient and do not form part of a particular communal game. The phrase “cool story, bro,” for example, is a meme that is quite widely used (across platforms and over time) as a sarcastic response to a story that is deemed pointless or boring.<sup>6</sup> In (63–65), the posters have affixed this response as a hashtag to their own tweets:

- (63) I just found out my Uncle’s ex-wife is the niece of Doris Roberts. #coolstorybro  
 (64) I woke up this morning in a panic because I thought it was Monday. #coolstorybro  
 (65) when I was younger I actually had an imaginary friend called Ralph.  
 #coolstorybro

The tag functions partly as a disclaiming meta-comment, acknowledging that the story related in the tweet is probably of little interest to anyone else and thus pre-empting a negative response from readers. Further, the tagging of the phrase “cool story, bro” can also be seen as way of recognizing its status and significance as an established cliché, i.e. meme, by marking it typographically.

Other memes are also treated in a similar manner, as in examples (66–68):

- (66) How is my Diet Pepsi flat inside the can? #idonteven  
 (67) Finally have a few hours for writing. Here I go. #nano #nevergonnagiveyouup#nevergonnaletyoudown  
 (68) Neighbors did not agree with our volume level. Denver PD was very cool about it though. #donttasemebro

These examples all have in common that the tagging functions to mark recognizable online memes typographically (in addition to other

---

<sup>6</sup> See e.g. <http://knowyourmeme.com/memes/cool-story-bro> (accessed 26 September, 2014).

communicative functions served by the tags in their respective contexts).<sup>7</sup> For instance, the exclamation “don’t tase me, bro!” was widely circulated online in a variety of contexts following the utterance of the words in a popular video clip of a young man who had to be forcibly subdued due to disruptive behavior during a speech by Senator John Kerry at the University of Florida. The poster of (68) is apparently repeating the meme as a joking contrast to her own peaceful encounter with Denver police.

Further searching also revealed a similar mode of hashtag usage for widely recognized quotes from popular films:

- (69) @user I had a run in today... Like a pack of wolves when they all join in.#franklymydearidontgiveadamn
- (70) Dear Knoxville, we are picking up trash tomorrow. Please drive carefully. Note, we are all stocked up on pumpkins.#GoAheadMakeMyDay
- (71) Robert DeNiro to host SNL on December 4th. #YouTalkinToMe ?
- (72) Did I mention that our pardon system has a 96% success rate, i.e. former inmates rehabilitated and NOT reoffending.#youcanhandlethetruth

In sequence, the tagged quotes come from *Gone With the Wind* (1939), *Sudden Impact* (1983), *Taxi Driver* (1976), and *A Few Good Men* (1992). Again, it seems likely that part of the reason why the different posters all chose to turn the references into hashtags might be to mark them typographically as significant clichés. “You talkin’ to me?”, for instance, is such a recognizable quote, even among many who have not seen the film, that it is essentially a cultural object in its own right.

#### 4. Discussion and conclusion

Hashtags on Twitter are used to perform a wide variety of communicative functions. For some uses, the linking feature of hashtags is directly relevant, but other times, it is clearly not. Some posters appear to be appropriating Twitter’s hashtag format as a substitute for features that Twitter lacks, e.g. tagging instead of bolding or italicizing. Other posters appear to be using tagging as an alternative to conventional options that Twitter does afford, e.g. marking emotive words and phrases with hashtags rather than with asterisks, or using hashtags instead of parentheses or

---

<sup>7</sup> All the memes in these examples are described in the meme database *Know Your Meme*. See e.g. <http://knowyourmeme.com/memes/dont-tase-me-bro> (accessed 26 September, 2014).

quotation marks. Many posters use hashtags in unpredictable ways, to mark punch lines or additional jokes in humorous tweets, or to mark Internet memes and pop culture references. In many cases, it is difficult to even speculate as to why the poster may have chosen to turn a word or phrase into a tag, but often the apparent arbitrariness of the tagging itself seems to carry playful force. Moreover, it seems plausible that even when hashtagging serves no other clearly discernible purpose, it can still serve some social purpose in establishing the poster's credentials as a member of the Twitter language community as opposed to other online communities.

Further, many instances of hashtag usage exemplified above seem notable for how they compress a lot of illocutionary force into short strings of letters, even if this happens at the expense of clarity. Previous CMC research has demonstrated an interest in the relation between affordances and constraints of new media and the forms of interaction and expression of language that take place in them (e.g. Thurlow & Brown 2003; Johnsen 2007; Spagnolli & Gamberini 2007; Crystal 2008; Tagg 2011). In a medium that imposes restrictions on utterance length, as Twitter does, it is possible that this is part of the reason why some posters put some content in a syntactically compressed hashtag form when it could otherwise have been typed out in full. However, it should also be noted that some of the tags analyzed above display playful excess and redundancy rather than compression.

As for the pragmatic approach taken here, David Crystal has suggested that a new approach to pragmatics might be necessary for dealing with CMC, since "classical" pragmatics is adapted to face-to-face speech situations (Crystal 2010: 234). However, the basic pragmatic approach of trying to unveil the logic by which intention is inferred from natural language utterances was definitely found to be applicable to the language situation of Twitter in this study. The meta-comment tags may be understood in terms of hedging, disclaiming and managing face, through the exploitation or flouting of maxims. The parenthetical explanations are analyzed as providing background information which is sometimes crucial to clarifying utterance force, but other times supplemental. The emotive and emphatic tags are analyzed mostly in terms of how they strengthen or change the illocutionary force of utterances, often in a manner reminiscent of the work done by non-verbal cues in face-to-face conversation. The humorous and playful uses of hashtags can be understood in terms of maxim-flouting and the exploitation of background knowledge. The addition of hashtags referencing memes and popular culture, of course, is

only made meaningful through inferences that the reader draws made on shared background knowledge of the origin of what is being referenced. While pragmatic methodology is typically applied to spoken interaction, this study demonstrates that even a traditional speech acts framework is eminently applicable to written communication in new media.

In sum, the findings reveal that users of Twitter have to a significant extent appropriated the hashtag organizing and categorizing device for other purposes, sometimes completely removed from the expected functionality. This may be a result of users actively extracting potential from a technology, but it is also possible to frame these new meanings and uses as affordances offered by the technology, or as functions emerging in the interface between user intentions and medium constraints and affordances (cf. Hutchby 2001; Mischaud 2007). The present study leaves this question open, but suggests it as an interesting avenue for further research into communication on Twitter. That being said, it can be concluded from the results presented above that users of Twitter have taken to the hashtag function, turning it into or tapping its potential as a multifunctional linguistic device for structuring information, playing games, and creating meaning in interaction.

## References

- Austin, John Langshaw; Urmson, James Opie (ed.) & Sbisà, Marina (ed.) (1976) *How to Do Things with Words: The William James Lectures Delivered at Harvard University in 1955*, second edition. Oxford: Oxford University Press.
- Baron, Naomi Susan (2008) *Always on: Language in an Online and Mobile World*. Oxford: Oxford University Press.
- (2009) The myth of impoverished signal: Dispelling the spoken language fallacy for emoticons in online communication. In Jane Vincent & Leopoldina Fortunati (eds.), *Electronic Emotion: The Mediation of Emotion via Information and Communication Technologies*, pp. 107–131. London: Peter Lang.
- Basch, Diego (2012, July 31) Some fresh Twitter stats (as of July 2012, dataset included). <<http://diegobasch.com/some-fresh-twitter-stats-as-of-july-2012>> (26 September 2014).
- boyd, danah & Ellison, Nicole (2008) Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication* 13 (1): 210–230.
- Crystal, David (2006) *Language and the Internet*, second edition. Cambridge: Cambridge University Press.
- (2008) *Txtng: The gr8 deb8*. Oxford: Oxford University Press.
- (2010) Internet language. In Louise Cummings (ed.), *The Pragmatics Encyclopedia*, pp. 234–236. London: Routledge.

- Dawkins, Richard (2006) *The Selfish Gene*, thirtieth anniversary edition. Oxford: Oxford University Press.
- Dresner, Eli & Herring, Susan (2010) Functions of the nonverbal in CMC: Emoticons and illocutionary force. *Communication Theory* 20 (3): 249–268.
- Deumert, Ana (2014) The performance of a ludic self on social network(ing) sites. In Philip Seargeant & Caroline Tagg (eds.), *The Language of Social Media: Identity and Community on the Internet*, pp. 23–45. Basingstoke: Palgrave Macmillan.
- Grice, Herbert Paul (1975) Logic and conversation. In Peter Cole & Jerry Morgan (eds.), *Syntax and Semantics: Vol. 3, Speech Acts*, pp. 41–58. New York: Seminar Press.
- Honeycutt, Courtenay & Herring, Susan C. (2009) Beyond microblogging: Conversation and collaboration via Twitter. In *Proceedings of the 42nd Hawaii International Conference on System Sciences*, pp. 1–10. Los Alamitos, CA: IEEE Press.
- Honigman, Brian (2012, November 11) 100 fascinating social media statistics and figures from 2012. *The Huffington Post*. <[http://www.huffingtonpost.com/brian-honigman/100-fascinating-social-me\\_b\\_2185281.html](http://www.huffingtonpost.com/brian-honigman/100-fascinating-social-me_b_2185281.html)> (26 September 2014).
- Huang, Jeff; Thornton, Katherine & Efthimiadis, Efthimis (2010) Conversational tagging in Twitter. In *Proceedings of the 21<sup>st</sup> ACM Conference on Hypertext and Hypermedia*, pp. 173–178. Toronto: ACM.
- Hutchby, Ian (2001) *Conversation and Technology: From the Telephone to the Internet*. Cambridge: Polity.
- Jansen, Bernard; Zhang, Mimi; Sobel, Kate & Chowdury, Abdur (2009) Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology* 60 (11): 2169–2188.
- Johnsen, Jan-Are (2007) Constraints on message size in quasi-synchronous computer mediated communication: Effect on self-concept accessibility. *Computers in Human Behavior* 23 (5): 2269–2284.
- Jonsson, Ewa (2013) *Conversational Writing: A Multidimensional Study of Synchronous and Supersynchronous Computer-Mediated Communication*. Uppsala University: Engelska institutionen.
- Kwak, Haewoon; Lee, Changhyun; Park, Hosung & Moon, Sue (2010) What is Twitter, a social network or a news media? In *Proceedings of the 19<sup>th</sup> International Conference on World Wide Web*, pp. 591–600. Raleigh, NC: ACM.
- Leech, Geoffrey (1983) *Principles of Pragmatics*. London: Longman.
- Levinson, Stephen (1983) *Pragmatics*. Cambridge: Cambridge University Press.
- Mischaud, Edward (2007) Twitter: Expressions of the whole self. An investigation into user appropriation of a web-based communications platform. London: Media@lse. <[http://www2.lse.ac.uk/media@lse/research/mediaWorkingPapers/MScDissertationSeries/Mishaud\\_Final.pdf](http://www2.lse.ac.uk/media@lse/research/mediaWorkingPapers/MScDissertationSeries/Mishaud_Final.pdf)> (26 September 2014).
- Murthy, Dhiraj (2011) Twitter: Microphone for the masses? *Media, Culture & Society* 33 (5): 779–789.
- (2013) *Twitter: Social Communication in the Twitter Age*. Cambridge: Polity.
- O'Connor, Brendan; Balasubramanian, Ramnath; Routledge, Bryan & Smith, Noah (2010) From tweets to polls: Linking text sentiment to public opinion time series.

- In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*. Washington, DC.
- Page, Ruth (2012) The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags. *Discourse & Communication* 6 (2): 181–201.
- Seargeant, Philip & Tagg, Caroline (2014) Introduction: The language of social media. In Philip Seargeant & Caroline Tagg (eds.), *The Language of Social Media: Identity and Community on the Internet*, pp. 1–20. Basingstoke: Palgrave Macmillan.
- Searle, John (1969) *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- (1979) *Expressions and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
- Spagnolli, Anna & Gamberini, Luciano (2007) Interacting via SMS: Practices of social closeness and reciprocation. *British Journal of Social Psychology* 46 (2): 343–364.
- Sperber, Dan & Wilson, Deirdre (1995) *Relevance: Communication and Cognition*, second edition. Oxford: Blackwell.
- Tagg, Caroline (2011) *The Discourse of Text Messaging: Analysis of Text Message Communication*. London: Continuum.
- Thurlow, Crispin & Brown, Alex (2003) Generation Txt? The sociolinguistics of young people's text-messaging. *Discourse Analysis Online* 1 (1): 30.
- Twitter (2014) Twitter terms of service. <<http://twitter.com/tos>> (26 September 2014).
- Weng, Jianshu; Lim, Ee-Peng; Jiang, Jing & He, Qi (2010) TwitterRank: Finding topic-sensitive influential twitterers. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, pp. 261–270. New York: ACM.
- Wu, Shaomei; Hofman, Jake; Mason, Winter & Watts, Duncan (2011) Who says what to whom on Twitter. In *Proceedings of the 20<sup>th</sup> World Wide Web Conference (WWW'11)*, pp. 705–714. Hyderabad, India: ACM
- Yus, Francisco (2011) *Cyberpragmatics: Internet-Mediated Communication in Context*. Amsterdam: John Benjamins.
- Zappavigna, Michele (2011) Ambient affiliation: A linguistic perspective on Twitter. *New Media Society* 13 (5): 788–806.
- (2012) *Discourse of Twitter and Social Media: How we Use Language to Create Affiliation on the Web*. London: Continuum.
- (2014) Enacting identity in microblogging through ambient affiliation. *Discourse & Communication* 8 (2): 209–228.

Contact information:

Peter Wikström  
 Karlstads universitet  
 Institutionen för språk, litteratur och interkultur  
 651 88 Karlstad, Sweden  
 e-mail: peter(dot)wikstrom(at)kau(dot)se

# **Squibs**

**María Luisa Carrió-Pastor**

## **Cross-cultural Variation in the Use of Modal Verbs in Academic English**

### **Abstract**

English academic writing has some specific characteristics that have been broadly defined by researchers. Nevertheless, English is undergoing constant modification as a result of being used as a lingua franca by international speakers. In this paper, my main objective is to determine whether language variation may be identified in cross-cultural communication when modal verbs of ability and possibility are used by speakers with different linguistic backgrounds. Furthermore, I would like to establish whether English writers tend to be more explicit than Spanish writers when both groups use English to communicate. The two corpora used in this study consisted of a set of fifty academic papers written in English by Spanish researchers and a set of fifty academic papers written in English by native English-speaking researchers. Both corpora were analysed to identify synchronic language variation in academic English when used by writers of different linguistic and cultural backgrounds. The results showed that there are disparities in the use of possibility and ability modal verbs and the conclusion reached is that writers with dissimilar mother tongues express volition through modal verbs differently in international journals.

### **1. Introduction**

The processes of second language writing are influenced by the writers' mother tongue, by linguistic conventions, and by genres, as researchers such as Charles (2007), Ozturk (2007), Samraj and Monk (2008), Durrant (2009), Hinkel (2002), Schlee (2009), Carrió-Pastor (2009, 2013), Carrió-

Pastor and Muñiz (2010), Carrió-Pastor and Candel Mora (2013) and Mauranen (2012, 2013a) have shown in their research. The aim of these analyses, which examine the use of a language by speakers with different mother tongues, is to demonstrate, through contrastive rhetoric, such as that which Kaplan (1966) and Connor (1996, 2002) have performed previously, that languages and writing traditions exhibit considerable synchronic variation.

Linguistic variation represents the different ways speakers of a language codify reality. In this vein, I believe human beings do not perceive concepts or express thoughts in the same manner. Consequently, their communication could vary according to their linguistic, cognitive, cultural and social background, following a functionalist approach. Thus, the *rules* for how language functions are not as general, as fixed or as evident as they are supposed to be when we study linguistic norms. Furthermore, if writers use an international language to communicate and it is not their mother tongue, their linguistic output may present variation due to the influence of their mother tongue.

In this study, I also consider the view that, although writers share the specialised knowledge of their field and are able to express their thoughts in a manner appropriate to an academic setting, they tend to use some processes from their mother tongues. As Duszak (1997: 9) points out, “Recent insights into academic writing have shown considerable variation in text characteristics across fields, languages and cultures. [...] Among the most notable differences are field- and culture-bound disparities in global organization schemata of texts.” Since English is the predominant language for such communication, linguists have studied the challenges faced by non-native English speakers when writing academic English (Salager-Meyer 2011; Carrió-Pastor & Muñiz 2010; Huang 2010; Carrió-Pastor 2013; Mauranen 2012, 2013a, 2013b, 2013c). At the same time, language use reflects social, linguistic, cultural, educational and professional conventions. One example of these conventions can be seen in academic English, the specific area of research of this paper.

Academic English is nowadays considered the international language for communicating with researchers all around the world and this means that the mode of expression employed by researchers should possess shared linguistic features, including short sentences, domain specific vocabulary and simple and direct language structures (Carrió-Pastor 2005, 2008, 2013; Wright 2008). In fact, writers tend to employ some of the cultural conventions of their own culture, with this conclusion having been reached

by researchers such as Hinds (1987, 1990), Connor and Mauranen (1999), Hyland (2005, 2008, 2010, 2011), Qi and Liu (2007) and Salager-Meyer (2011). These authors make use of the notion of reader responsibility and writer responsibility, with the allocation of this responsibility varying according to the culture of the writer. They maintain that, in some languages, writers are accountable for effective communication, whereas there are languages in which readers are the ones responsible for understanding a text. These researchers consider that writers may be liable for providing landmarks or transition statements in the text. These transition devices can be explicit and so it is the writer who must lead the way for the reader; or they can be implicit and may require an active role from the reader, who must use his or her intuition to follow the text. Modal verbs can act as such landmarks, transmitting information to the readers and also the opinion of the writer. The existence of these landmarks could have implications for the teaching of academic writing, as well as for cross-cultural understanding between academics.

In addition, here I presume that the fact that the globalization of information and the Internet have changed cross-cultural communication. Intercultural communicative competence can vary depending on the cultures of the speakers involved and, as a result, speakers with different cultural backgrounds may not conceive the relationship that exists between concepts and words in the same way, as researchers such as Canagarajah (2002, 2007a, 2007b), Arasaratnam and Doerfel (2005), Zhou (2008), Pennycook (2010) and Louhiala-Salminen and Kankaanranta (2012) have explained. This is one of the main reasons why there can be different ways in which to transmit the same reality in a language used in cross-cultural communication; a strict standard production or interpretation of language is not advisable in an era in which change and development are key features of it. This era also sees increased interaction between people from different cultures and languages: a monocultural perspective on language use is no longer acceptable.

In this sense, Kramsch (1998: 3) states that “Speakers identify themselves and others through their use of language; they view their language as a symbol of their social identity [...] Thus we can say that language symbolizes cultural reality.” Through language, speakers transmit their own perception of reality and they use it to persuade, influence or manipulate their audience, as considered by Mauranen, Hynninen and Ranta (2010) in the ELFA project. Speakers’ choices of linguistic elements

and their use of specific rhetorical items may reflect how they perceive the world and how they wish to transmit information and ideas.

In this study, I decided to focus on the collection of examples and analysis of the modal verbs *can*, *could*, *may*, *might*, and *be able to*, as they have a range of pragmatic functions and serve to mark evidentiality, possibility and likelihood, strategic vagueness, and politeness in discourse (Channell 1994; Markkanen & Schroeder 1997; Alonso Almeida 2012). As they express modality, their use is greatly influenced by the mother tongue of the writer, as researchers such as Hinkel (2009), for example, have shown. The writer transmits cultural and social conventions through modal verbs and yet Hinkel (2009: 672), when discussing the pragmatic properties of modals in academic writing, has noted the following: “Comparatively fewer studies have addressed the uses of these modals specifically in student L1 and L2 writing, with a likely exception of their uses as hedges, qualifiers, or markers of (un)certainty”. This is the reason why I believe modal verbs need to be analysed, while also taking into account the implications of their use by speakers with different mother tongues.

In order to explore all the aspects mentioned above, the main objective of this paper is to determine the way in which possibility and ability modal verbs are used by native speakers of English and non-native speakers of English in an academic context. More specifically, my intention is to identify whether language variation may be identified in cross-cultural communication when modal verbs of ability and possibility are used by speakers with different linguistic backgrounds. Finally, an attempt is made to establish whether native English-speaking writers tend to be more explicit than Spanish writers who use English as a foreign language to communicate internationally.

## **2. Methodology**

The two corpora used in this study consisted of, on the one hand, a set of fifty academic papers written in English by researchers from Spain and which were published in international journals from 2010 to 2012 (referred to as the NNES corpus hereinafter). On the other hand, I also compiled a second corpus composed of fifty papers written by native English-speaking researchers and published in international journals from 2010 to 2012 (referred to as the NES corpus hereinafter). The total number of words included in the NNES corpus was 184,357 (47.11% of the total corpus) and

the total number of words included in the NES corpus was 206,907 (52.89% of the total corpus).

The papers included in both corpora were selected from the subject domain of engineering and, afterwards, classified by considering the name and affiliation of the author or, in case of multiple authorships, the name and affiliation of at least two or three authors, depending on the number of authors involved in the research. The criteria on which the articles comprising each corpus were selected included authorship, length, the nature of the texts (academic English) and the likely audience (international academic community).

Once the papers were compiled, all the tables, direct citations, graphs, charts, bibliography and references were removed and the corpora were analysed using the *Wordsmith Tools* suite of computer programs, version 5.0 (Scott 2009). The modal verbs to be analysed were those that indicated possibility and ability and *can, could, may, might and be able to* were chosen because they are identified as verbal modality markers and they have a range of textual and pragmatic functions. Later, these modal verbs were identified in the corpora. In this study, I carried out a quantitative analysis of the corpora in order to analyse the variation in the use of modal verbs, but a qualitative analysis of the cases found was also performed in order to determine in context whether the English writers tended to be more explicit than the Spanish writers.

One program from the Wordsmith Tools suite, the *Concordancer*, proved particularly useful in the identification of modal verb variations in the corpora. Nevertheless, manual checking and identification was also necessary in order to confirm the use of some patterns in the academic papers. I identified the occurrences found in the corpora and classified them into the different modal verbs, calculating the percentages and the statistical data. The chi-squared value was calculated, with it being necessary for the p-value to be lower than 0.05 for the results to be statistically significant. The relative risk was also calculated in order to establish the probability of this value appearing in similar analyses, establishing 1 as the minimum value.

After compiling the results for the modal verbs of possibility and ability, comparisons were drawn between the corpus of non-native English speakers (NNES) and the corpus of native English speakers (NES) in order to observe whether variation could be detected in the use of these verbs. I also analysed and described some examples of the modals found in both corpora in order to observe the degree of judgment and vagueness

associated with their use. Finally, the results were analysed, the most interesting data were identified, and the conclusions of the study were drawn.

### 3. Results and discussion

The results showed that there is some variation in the use of possibility and ability modals by writers with different linguistic and cultural backgrounds, namely between Spanish and English writers in this case. Although these writers share the specialised knowledge and express their thoughts in an academically appropriate way, these verbs qualify the meaning of the utterance made by the writer. The possibility and ability modals may express the judgement of the writer in a nuanced way and express different gradients of possibility and may be used with another meaning in different parts of the sentence, depending also on the intention and the cultural tradition of the speaker.

The occurrences found in the corpora of the use of possibility and ability modal verbs can be observed in Table 1. The second and fourth columns show the results expressed in percentages taking into account the total occurrences of each modal verb. The data of the third and fifth columns are calculated taking into consideration the total amount of words of the NNES corpus and the NES corpus.

**Table 1.** Frequencies of modal verbs used by NNES and NES

Modal verbs	NNES corpus percentage	NNES per 1,000 words	NES corpus percentage	NES per 1,000 words
CAN	<b>59.82</b>	<b>4.75</b>	40.18	<b>2.84</b>
COULD	48.82	0.90	51.18	0.84
MAY	39.69	0.98	<b>60.31</b>	<b>1.32</b>
MIGHT	24.07	0.07	<b>75.93</b>	0.19
BE ABLE TO	<b>76.47</b>	0.42	23.53	0.11

The results show that NES used the modal verbs *may* and *might* more frequently (in bold), with almost 61% and 76% of the occurrences respectively; while the NNES tend to rely on the modal verb *can*, with almost 60% of the occurrences. The results presented in Table 1 support the observation that the manner of being imprecise and appearing polite in formal writing is sometimes determined by cultural and social conventions, as there is a difference in the use of modal verbs of possibility and ability. Writers with different linguistic backgrounds may not share the same

cultural norms, and, when this happens, variation appears in modality. I observed in the examples that Spanish writers tended to express possibility with the verb *can* (*poder* in Spanish) and *could* (*podría* in Spanish) was used to express weak possibility, not politeness as happened in the corpus of NES. This could be due to the linguistic reference to the mother tongue of the Spanish writers. We can see some examples of the use of *can* by NNES and NES in (1) and (2):

(1) “For a metallic electrode, the attenuation *can* be calculated using its conductivity and the values of  $n_0$  and  $Z$  in an analytical formula”. “It is found that a substantial improvement in bandwidth *can* be obtained, provided that the velocity of the optical and electrical signals is matched”. (NNES)

(2) “As *can* be appreciated in Fig. 1, the XRD of the YBCO/YSZ/LNO heterostructure reveals that, after the YBCO growth, the YSZ layer suffers a significant crystalline deterioration, and crystals with orientation are observed”. (NES)

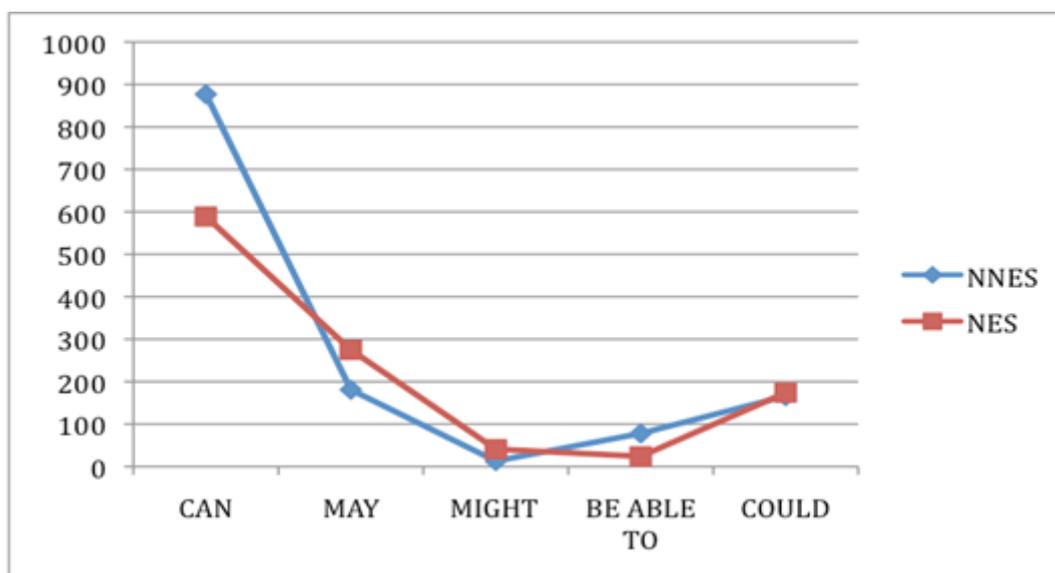
In (1), the modal verb *can* could be substituted by a non-assertive modal such as *may* or *might*, but the NNES preferred to use *can*, instead. This may mean that Spanish writers prefer to be precise and transmit a judgement, not considering the possibility that the reader may interpret possibility as assertion. The translation into Spanish of this part of the sentence is: “Se ha visto que una mejora substancial en el ancho se banda se puede obtener...”, with *puede* being the only possible verb in this context in Spanish as the conditional *podría* (*may; might*) should not be used, as it may transmit the idea of conditionality or improbability to the reader. In (2), on the other hand, the modal verb *can* cannot be substituted by another modal verb, as the writer expresses what the reader can see in a table, so the writer is expressing an objective reality and ability.

The statistical analysis of the results can be observed in Table 2. The second column shows the calculation of the relative risk of the results obtained in both the NNES and NES corpora. As the minimum value established is 1, the results obtained in this analysis can be extrapolated to similar research. The third column illustrates the results of the p-value, which significance level is 0.05. As all the values are below 0.05, it was obtained a very strong presumption against the neutral hypothesis, i.e. the results obtained are significant.

**Table 2.** Statistical analysis of data obtained

Modal verbs	Relative risk	Chi-squared
CAN	1.26 (1.16–1.36)	p= 0.00
COULD	0.81 (0.66–0.99)	p= 0.03
MAY	0.56 (0.47–0.66)	p= 0.00
MIGHT	0.27 (0.14–0.50)	p= 0.00
BE ABLE TO	2.75 (1.75–4.32)	p= 0.00

After this analysis, an overall comparison of the results obtained from the corpora was undertaken. Figure 1 shows that NNES used more possibility and ability modals, and it should also be noted that most of the difference found in the occurrences was related to the use of *can* by NNES.

**Figure 1.** Comparison of the use of possibility and ability modals in the corpora.

We can also see that NNES and NES used *be able to* and *might* quite similarly, but that there was divergence in the use of *may* and *can*. As observed, in formal academic writing, the concept of vagueness can be expressed in a different way by writers, depending on the cultural background of the authors. Although the meaning of an academic text should be appropriately qualified and reflect the opinion of the writer, sometimes writers make some changes in the way they express emotions or volition. A research paper should communicate its results and conclusions in an impartial manner, so as not to manipulate the reader, but objectivity is not easily maintained when writing about ideas and perceptions in another language than the mother tongue. As pointed out in the Introduction section, languages can be either reader or writer-responsible, with English being the latter. Academic English readers expect landmarks of modality as

they read, and writers need to provide them. Writers with different linguistic backgrounds communicate in English as a lingua franca transmitting their cultural conventions, which may enrich language production.

Although English-speaking readers expect writers of English to be explicit and direct, it is not always the case that they are. English is a global language and is used by speakers with different linguistic and cultural backgrounds. The need for imprecision or precision is determined by the specific context and writers must make their choice on which modal verb to use by taking the context into account, but it is also the case (as seen in the results of this research) that the level of precision expected in formal writing is sometimes governed by cultural conventions. As a consequence, readers and writers do not employ modal verbs of possibility and ability in the same way, as demonstrated by the results displayed in Table 1 and Graph 1.

#### 4. Conclusions

My purpose has been to demonstrate that linguistic and cultural backgrounds may influence the way in which non-native English writers express their ideas in a lingua franca and that this leads to variation in the use of the language that most international writers employ in order to explain scientific findings correctly, i.e. English. I consider that the difference between the linguistic and the cultural background is that a language expresses shared conventions among the speakers of the same language, but our cultural background is the factor that differentiates the way we express ourselves in a language and leads to the formulation of different ways of expressing the same reality. I believe that different cultural backgrounds cause variation in choice of expressions in a given language.

In addition, variations in the transmission of possibility and ability modals in international papers are not a weird phenomenon as speakers transmit more than words when they communicate their ideas fully in an environment of cross-cultural communication. Furthermore, I also believe that English as an international language may also reflect different ways of expressing the same reality and readers should be conscious of this fact. As an example, the data seen in Table 1 and Figure 1 show that NES express possibility mainly with *can*, *may* and *might*, meanwhile NNES prefer the use of *can*, as, in some contexts in Spanish, the equivalent of *may* and

*might* expresses a weaker possibility (e.g. *podría ser* could be interpreted as not being real, as it expresses improbability).

I think that due consideration should be given to language variation, as the connection between the conventions of culture and language is not as straightforward as we might think. Variation also exists in the use of language among native speakers and so, logically, variation may appear when non-native speakers use an international language. Given that there are as many ways of transmitting information as there are speakers of a language, to think that there is one standard form of that language in an international setting seems too idealistic and unsuited to the needs of international communication.

If we consider that reality can be represented or transmitted in different ways, then it is quite logical to think that second language speakers do not communicate in the same way as native speakers, as demonstrated by the research carried out by Charles (2007), Ozturk (2007), Samraj and Monk (2008), Durrant (2009), Hinkel (2009), Schlee (2009), Carrió-Pastor (2009, 2013), Carrió-Pastor and Muñiz (2010), Carrió-Pastor and Candel (2013) and Mauranen (2012, 2013a, 2013b, 2013c). Although second language writers possess the same level of language proficiency, when these authors communicate in an international context, some differences can be detected which are related to their native cultural and social conventions.

The results of this study of modal verbs of possibility and ability show that NNES made greater use of *can*, which expresses the possibility of the proposal. In this sense, Spanish writers of English as a second language tend to transmit less vagueness when they express themselves in English than their native-speaking English counterparts. This seems to indicate that the cultural background of the writer influences the style of writing, and, hence, the writer's standpoint. The style of the writer might sometimes appear to lack objectivity and be a reflection of the writer's judgement of the likelihood of the truthfulness of a particular proposition; nevertheless, individual ways of communication should be accepted. They enrich the language and reflect the changes it suffers as it is being used by millions of speakers. It is a part of its natural evolution. In the corpus analysed in this study, I have observed that Spanish writers tend to be more assertive and this fact has been noted in previous studies (Carrió-Pastor 2005, 2009). This may impact upon the production of language and the different ways of transmitting ideas.

The variation observed could be attributed to the different rhetorical and educational traditions in academic writing in English and Spanish, but it can also form the basis for an appeal for a better understanding and tolerance of culture-specific features, with a view to preserving cultural identity when using English as the international language of academic communication, as Vold (2006) and Vassileva (2001) have stated in their research. Since language constitutes the vehicle for the transmission of thought, it may be difficult to establish guidelines with which to detect modal variation. In future studies, this aspect should be taken into account. First, it might be necessary to determine which parts of the text are the transmissions of modality and differentiate them from those that derive from the application of language rules. Second, the process followed by the speaker in order to decide on the modal verb to be used should be clearly established and, third, other possible ways to express modality should be taken into account by the writer.

The results of this study lead us to the view that synchronic variation in academic discourse is most probably caused by the influence of the cultural background of the writers. As we have seen, the importance of this topic has recently been reflected in the literature as shown in section 1, and it is my view that the different use and variation of possibility and ability modals as hedges and politeness devices should be included as part of English academic writing courses – this is of interest to international speakers of English. At the same time, it can be stated that academic English is sensitive to change in cross-cultural communication and may show greater variability, given that it is widely used for communication all over the world.

## References

- Alonso Almeida, Francisco (2012) An analysis of hedging in eighteenth century English astronomy texts. In Isabel Moskowich & Begoña Crespo (eds.), *Astronomy “playne and simple”: The Writing of Science between 1700 and 1900*, pp. 199–220. Amsterdam: John Benjamins.
- Arasarrantnam, Lily A. & Doerfel, Marya L. (2005) Intercultural communication competence: Identifying key components from multicultural perspectives. *International Journal of Intercultural Relations* 29: 137–163.
- Canagarajah, A. Suresh (2002) *Critical Academic Writing and Multilingual Students*. Ann Arbor: University of Michigan Press.
- (2007a) The ecology of global English. *International Multilingual Research Journal* 1(2): 89–100.

- (2007b) Lingua franca English, multilingual communities, and language acquisition. *Modern Language Journal* 91(5): 921–937.
- Carrió-Pastor, María Luisa (2005) *Contrastive Analysis of Scientific-technical Discourse: Common Writing Errors and Variations in the Use of English as a Non-native Language*. Ann Arbor: UMI.
- (2008) English Complex Noun Phrase Interpretation by Spanish Learners. *Revista Española de Lingüística Aplicada* 21: 27–44.
- (2009) Contrasting specific English corpora: Language variation. *International Journal of English Studies* 14(1): 221–234.
- (2013) A contrastive study of the variation of sentence connectors in academic English. *Journal of English for Academic Purposes* 12(3): 192–202.
- Carrió-Pastor, María Luisa & Muñiz, Rut (2010) Variations in business English letters written by non-native writers. *LFE. Revista de Lenguas para Fines Específicos* 16: 39–56.
- (2012) Lexical variations in business e-mails written by non-native speakers of English. *LSP Professional Communication, Knowledge Management and Cognition* 3(1): 4–13.
- Carrió-Pastor, María Luisa & Candel Mora, Miguel Ángel (2013) Variation in the translation patterns of English complex noun phrases into Spanish in a specific domain. *Languages in Contrast* 13(1): 28–45.
- Connor, Ulla (1996) *Contrastive Rhetoric: Cross-Cultural Aspects of Second Language Writing*. Cambridge: Cambridge University Press.
- (2002) New directions in contrastive rhetoric. *TESOL Quarterly* (36)4: 493–510.
- Connor, Ulla & Mauranen, Anna (1999) Linguistic analysis of grant proposals: European Union research grants. *English for Specific Purposes* 18: 47–62.
- Channell, Joanna (1994) *Vague Language*. Oxford: Oxford University Press.
- Charles, Maggie (2007) Argument or evidence? Disciplinary variation in the use of the noun that pattern in stance construction. *English for Specific Purposes* 26: 203–218.
- Durrant, Paul (2009) Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes* 28: 157–169.
- Duszak, Anna (1997) Cross-cultural academic communication. In Anna Duszak (ed.), *Culture and Styles of Academic Writing*, pp. 11–40. New York: Mouton de Gruyter.
- Hinds, John (1987) Reader versus writer responsibility: A new typology. In Ulla Connor & Robert Kaplan (eds.), *Writing Across Languages: Analysis of L2 Text*, pp. 141–152. Reading: Addison-Wesley.
- (1990) Inductive, deductive, quasi-inductive: Expository writing in Japanese, Korean, Chinese, and Thai. In Ulla Connor & Anne M. Johns (eds.), *Coherence in Writing: Research and Pedagogical Perspectives*, pp. 87–110. Alexandria, VA: TESOL.
- Hinkel, Eli (2002) *Second Language Writers' Text: Linguistic and Rhetorical Features*. New Jersey: Lawrence Erlbaum Associate.
- (2009) The effects of essay topics on modal verb uses in L1 and L2 academic writing. *Journal of Pragmatics* 41: 667–683.

- Huang, Ju Chuan (2010) Publishing and learning writing for publication in English: Perspectives of NNES PhD students in science. *Journal of English for Academic Purposes* 9: 33–44.
- Hyland, Ken (2005) Representing readers in writing: Student and expert practices. *Linguistics and Education* 16: 363–377.
- (2008) ‘Small bits of textual material’: A discourse analysis of Swales’ writing. *English for Specific Purposes* 27: 143–160.
- (2010) Constructing proximity: Relating to readers in popular and professional science. *Journal of English for Academic Purposes* 9: 116–127.
- (2011) Projecting an academic identity in some reflective genres. *Ibérica* 21: 9–30.
- Kaplan, Robert B. (1966) Cultural thought patterns in intercultural education. *Language Learning* 16: 1–20.
- Kramsch, Claire (1998) *Language and Culture*. Oxford: Oxford University Press.
- Louhiala-Salminen, Leena & Kankaanranta, Anne (2012) Language as an issue in international internal communication: English or local language? If English, what English? *Public Relations Review* 38: 262–269.
- Markkanen, Raija & Schroeder, Hartmut (1997) Hedging: a challenge for pragmatics and discourse analysis. In Raija Markkanen & Hartmut Schroeder (eds.), *Hedging and Discourse: Approaches to the Analysis of a Pragmatic Phenomenon in Academic Texts*, pp. 3–21. Berlin: Walter de Gruyter.
- Mauranen, Anna (2012) *Exploring ELF: Academic English Shaped by Non-native Speakers*. Cambridge: Cambridge University Press.
- (2013a) Lingua franca discourse in academic contexts: Shaped by complexity. In John Flowerdew (ed.), *Discourse in Context: Contemporary Applied Linguistics*, vol. 3, pp. 225–245. London: Bloomsbury Academic.
- (2013b) Speaking professionally in an L2 – Issues of corpus methodology. In Julia Bamford, Silvia Cavalieri & Giuliana Diani (eds.), *Variation and Change in Spoken and Written Discourse. Corpus Linguistic Approaches*, pp. 5–31. Amsterdam: John Benjamins.
- (2013c) ‘But then when I started to think...’ Narrative elements in conference presentations. In Maurizio Gotti & Carmen Sancho Guinda (eds.), *Narratives in Academic and Professional Genres*, pp. 45–66. Bern: Peter Lang.
- Mauranen, Anna; Hynninen, Niina & Ranta, Elina (2010) English as an academic lingua franca: The ELFA project. *English for Specific Purposes* 29: 183–190.
- Ozturk, Ismet (2007) The textual organisation of research article introductions in applied linguistics: Variability within a single discipline. *English for Specific Purposes* 26: 25–38.
- Pennycook, Alastair (2010) *Language as a Local Practice*. New York: Multilingual Matters.
- Qi, Xiukun & Liu, Lida (2007) Differences between reader/writer responsible languages reflected in EFL learners’ writings. *Intercultural Communication Sciences* XVI(3): 148–159.

- Salager-Meyer, Françoise (2011) Scientific discourse and contrastive linguistics: explicitness and the concept of reader/writer responsible languages. *European Science Editing* 37(3): 71–72.
- Samraj, Betty & Monk, Lenore (2008) The statement of purpose in graduate program applications: Genre structure and disciplinary variation. *English for Specific Purposes* 27: 193–211.
- Schleef, Erik (2009) A cross-cultural investigation of German and American academic style. *Journal of Pragmatics* 41: 1104–1124.
- Scott, Michael (2009) *Wordsmith Tools 5*. Liverpool: Lexical Analysis Software.
- Vassileva, Irena (2001) Commitment and detachment in English and Bulgarian academic writing. *English for Specific Purposes* 20(1): 83–102.
- Vold, Eva Thue (2006) Epistemic modality markers in research articles: a cross-linguistic and cross-disciplinary study. *International Journal of Applied Linguistics* 16(1): 61–87.
- Wright, Laura J. (2008) Writing science and objectification: Selecting, organizing, and decontextualizing knowledge. *Linguistics and Education* 19: 265–293.
- Zhou, Yan (2008) The impact of cultural transfer on cross-cultural communication. *Asian Social Science* 7(7): 142–146.

Contact information:

María Luisa Carrió-Pastor  
Universitat Politècnica de València  
Departamento de Lingüística Aplicada  
Edificio 4P, Camino de Vera, 14  
46022 Valencia (Spain)  
e-mail: lcarrio(at)upv(dot)es

Maurer, Philippe (2009) *Principense (Lung'le). Grammar, Texts, and Vocabulary of the Afro-Portuguese Creole of the Island of Príncipe, Gulf of Guinea*. London: Battlebridge Publications. Pp. viii + 280.

Reviewed by Angela Bartens

The book under review is a comprehensive description of the highly endangered Portuguese-lexifier creole of Príncipe – henceforth Principense –, formerly a Portuguese colony and since 1975 part of the independent Republic of São Tomé and Príncipe. So far, scholars have essentially had to rely on Günther (1973) for information on Principense. Maurer's monograph turns it into one of the best described creoles, especially when considering the size of the speech community which has less than a hundred speakers, a core group being constituted by elderly women. The language is under siege by Portuguese and Sãotomense or Forro, the main creole of the sister island São Tomé. The author attributes the obsolescence of the language to several factors (pp. 3–4). Firstly, a sleeping sickness epidemic befell the island around 1900. According to the sources consulted by Günther (1973: 12–13), only approximately 300 inhabitants of Príncipe survived. By consequence, indentured laborers from Angola, Mozambique, and the Cape Verde Islands had to be imported. For three or four generations, Principense has no longer been passed on to younger people nor have efforts been made to give speakers of the Cape Verdean Creoles or other languages access to Principense. What we find thence is a territorially heterogeneous speech community of multilingual speakers whose main means of daily communication is no longer constituted by Principense. Their linguistic production features the influence of language contact to varying degrees and it seems that both linguistic and meta-linguistic competence is really fading as, for instance, Maurer's informants were able to indicate only a few lexical and morphological isoglosses which may reflect former dialectal variants (p. 4). Therefore the volume at hand is even more valuable to creolists and anyone interested in the vanishing manifestations of linguistic diversity in that particular region of the world – and beyond.

The book is organized as follows: a short introductory chapter (pp. 1–6) gives the reader basic information about the sociohistorical and sociolinguistic setting; it also introduces previous work on Principense as

well as the corpus and the way in which examples are presented. The description of the language structure is divided into phonology (pp. 7–28), morphosyntax (pp. 29–172), and “Miscellaneous” (pp. 173–178). This section which I will comment upon below is followed by an anthology of texts (pp. 179–220; the shorter texts are both glossed and translated, the longer ones only translated), a comprehensive Principense-English word list of 1650 entries (pp. 211–244), and a shorter one into the reverse direction (pp. 245–256). The original Principense story “The mouth that says good things also says bad things” has been translated into the two other creole languages spoken in the Republic of São Tomé and Príncipe, Lung’Ie and Lunga Ngola, in order to give the reader an idea of the similarities as well as the differences between these languages. This task is facilitated by a comparative word list. The parallel texts plus the comparative word list constitute Appendix I (pp. 257–260) whereas Appendix II (pp. 261–274) is constituted by a critical edition of Ribeiro (1888), the first known grammar of Principense. The inclusion of this critical edition as well as the fact that several audio files of materials presented in the book can be freely downloaded from the editorial’s website ([www.battlebridge.com/books/maurer/audio\\_files.html](http://www.battlebridge.com/books/maurer/audio_files.html)) greatly enhance this work which is, as I have already pointed out, extremely valuable for its description of the language structure of Principense. The volume also contains a table of contents (pp. iii–iv), acknowledgements (p. v), a list of references (pp. 275–276), and an index (pp. 277–280).

After commenting briefly on some aspects of the descriptions of the phonological system and the morphosyntax of Principense, I shall devote the remainder of this review to the chapter dealing with “miscellaneous” language features. For example, Maurer presents a new analysis of the tonal system of Principense, arguing that the language has two tones (H, L) with sequences of tones but no contour tones (p. 14–26). The phonological status of nasal vowels and implosive and labiovelar stops appears to be unclear as far as the present phonological system is concerned (cf. pp. 8–9), suggesting that some information on the original language structure may have been lost for good. Likewise, the validator *na* (pp. 67–68) was not identified by Günther (1973), the main previous scholarly work. Whereas Günther’s observations on sentence-final particles (1973: 171) formulated in a telegraphic style attribute clear functions to each of them, Maurer’s description of their use reveals significant overlap (pp. 169–170). It might be hypothesized that this be, at least in part, due to linguistic insecurity in the speakers. On the other hand, it becomes quite clear from the examples

given, for instance, for the use of *ê*, that the functions of vocative, interrogative and validating *ê* are actually interrelated.

As for the features listed by Maurer as miscellaneous, two phenomena strike me as particularly worthy of commenting. Firstly, the author discusses reduplication, identifying the following functions: intensification, indefiniteness, and distributiveness (with numerals). Reduplication is also involved in the substrate-driven formation of numeral plurals as exemplified by *dexi dexi kumin dôsu* ‘ten ten place two = twenty’ and *sen sen kumin têêxi* ‘hundred hundred place three = three hundred’ (p. 174). As for the second phenomenon, Maurer observes that in two (and only two – nevertheless, there may be others not present in his data) cases, “the partial or complete reduplication has the effect of changing the syntactic category.” The examples he gives, *kêtê* ‘small’ vs. *kêtê-kêtê* ‘little by little’ and *moli* ‘soft’ vs. *moli-moli* ‘slowly’, actually reflect Euro-centric divisions into word classes which are frequently not retained in creole languages – nor do they exist cross-linguistically (cf. Haspelmath 2007). Ladhams et al. (2003: 173) find that Principense patterns with its main substrate Edo in using reduplication to express intensification and distributiveness. In the terminology used by Bartens (2004: 239, 241, 244), this language also reduplicates nouns to express accumulation and verbs to express duration.<sup>1</sup> Indeed, for instance Maurer’s example (1250) of intensifying reduplication, *Mene sa udêntu matu a vya vya...* ‘Mene was in the forest turning around...’ (p. 174), actually conveys the duration of the action of turning around. By consequence, we have to conclude that the categories adopted for the classification of linguistic phenomena may have quite far-reaching effects on their subsequent analyses. The category of ‘indefiniteness’, also present in the taxonomy of Bartens (2004) as ‘indeterminacy’, was not identified in the Principense materials scrutinized for that study. It is interesting to note that all the reduplicated elements in Maurer’s category ‘indefiniteness’ are nouns which alone or modified by the determiner *ki* ‘what’ function as interrogative pronouns: *kumin* ‘place’ vs. *kumin kumin* ‘wherever’, *kwa* ‘thing’ vs. *kwa kwa* ‘some’, *modi*

---

<sup>1</sup> Ladhams et al. (2003: 167) state with regard to “intensive noun reduplication” that “[w]e use the term ‘intensive’ here to refer, first, to noun reduplication with a cumulative effect.” The possibility of reduplication expressing the duration of a verbal action in both Lunga Ngola and Lung’le is likewise mentioned in Ladhams et al. (2003: 169), in both cases based on Maurer’s data. Neither observation is repeated in the volume under review.

‘manner’ vs. *modi modi* ‘be it as it were’, *ningê* ‘person’ vs. *ningê ningê* ‘whoever; whomever’ (pp. 45, 174).

Whereas ideophones are usually described as items which prototypically modify only one lexical item (cf. Westermann 1907: 83), Maurer shows in his list of ideophones (p. 176) that some ideophones actually modify two (or more) items. His observation according to which “[i]t seems that ideophones that modify adjectives which refer to permanent states [...] are more likely to occur without the modified word” is a hypothesis intuitively worth exploring in other creoles as well. This hypothesis also allows for the postulate of ideophones occurring without (overtly) modifying another lexeme (cf. Bartens 2000: 42).

Summarizing, Maurer’s work is a valuable resource for creolists and other scholars and a landmark in the study of creoles in general and the Portuguese-lexifier creoles of the Gulf of Guinea in particular.

## References

- Bartens, Angela (2000) *Ideophones and Sound Symbolism in Atlantic Creoles*. Helsinki: The Finnish Academy of Science and Letters.
- Bartens, Angela (2004) A comparative study of reduplication in Portuguese- and Spanish-based creoles. In Mauro Fernández, Manuel Fernández-Ferreira & Nancy Vázquez Veiga (eds.), *Actas del III Encuentro de ACBLPE*, pp. 239–253. Madrid: Iberoamericana & Frankfurt: Vervuert.
- Günther, Wilfried (1973) *Das portugiesische Kreolisch der Ilha do Príncipe*. Marburg an d. Lahn: author’s edition.
- Haspelmath, Martin (2007) Pre-established categories don’t exist: Consequences for language description and typology. *Linguistic Typology* 11: 119–132.
- Ladhams, John; Hagemeyer, Tjerk; Maurer, Philippe & Post, Marike (2003) Reduplication in the Gulf of Guinea languages. In Silvia Kouwenberg (ed.), *Twice as Meaningful. Reduplication in Pidgins, Creoles and Other Contact Languages*. Westminister Creolistics Series 8, pp. 165–176. London: Battlebridge.
- Ribeiro, Manuel Ferreiro (1888) *Dialecto da ilha do Príncipe*. MS no. 11.23.12, Schuchardt Archives of the University of Graz, Austria.
- Westermann, Diedrich (1907) *Grammatik der Ewe-Sprache*. Berlin: Dietrich Reimer.

Contact information:

Angela Bartens  
School of Languages and Translation Studies  
Faculty of Humanities  
FI-20014 University of Turku

Department of Modern Languages  
PO Box 24  
FI-00014 University of Helsinki  
e-mail: angbar(at)utu(dot)fi



Susana S. Fernández & Johan Falk (eds.) (2014) *Temas de gramática española para estudiantes universitarios, Una aproximación cognitiva y funcional*. Frankfurt am Main: Peter Lang. Pp. 317.

Reviewed by Ilpo Kempas

As expressed by its title, the compilation in question is aimed at university students of Spanish Philology. Most of its chapters are dedicated to topics related to verbal syntax, but also the foundations of cognitive grammar as well as the Spanish prepositions are focused on. The book is especially targeted at Nordic students, for which it adopts a contrastive approach in some of its chapters.

The book provides interesting details and insights along the way. Instead of presenting overwhelmingly new research data, however, the book focuses on explaining some of the key questions of Spanish grammar in a way easily intelligible to the student. In my view, this goal is achieved successfully, and the book would serve well as study material on courses related to the covered topics. Although the approach followed in the book is cognitive, all the figures are easily intelligible to readers without previous familiarity with the conventions of this field of linguistics. The pedagogical aims of the book can also be seen in exercise boxes related to each topic, in which the reader is invited to reflect on a certain aspect of each topic in relation to their native language.

The book is written by several experts working at universities in the Nordic Countries, with long experience in teaching Spanish to non-native speakers. Although written by several authors, the end result is a coherent work, with no thematic overlapping.

Chapter 1, written by Teresa Cadierno and Johan Pedersen, is a general introduction to cognitive linguistics. In my opinion, this complex area is resumed in a clear and illustrative way. The chapter starts by presenting language as a cognitive phenomenon (pp. 15–17), discussing next the origin and tenets of cognitive linguistics (pp. 17–19). The following pages (pp. 19–31) offer the reader a concise and relevant account of the key concepts of this field of linguistics (categorization, iconicity, indexicality, nature of meaning). This is followed by a presentation of how language acquisition and learning are viewed by cognitive linguistics (pp.

31–34). On page 32, the authors also point out the role of *construction* as the basic unit of language.

In Chapter 2, Uwe Kjær Nissen deals with prepositional accusative objects and dative objects. The author looks at them from the functional perspective, with *agent* and *patient* as the key concepts. Morphologically similar forms with different functions may easily cause ambiguity among learners of Spanish. This is the case of preposition *a*, the different semantic roles of which the author discusses. On page 46, the author takes a look at the different parameters of use of preposition *a*, based on the patient roles the preposition may take, illustrating a gradual shift from more prototypical to less prototypical as the degree of humanity of the patient increases. The following pages (pp. 47–58) are dedicated to the different uses of the Spanish dative. Interestingly, in the case of Spanish, the traditional concept of *dative* seems to serve to distinguish – at least on the practical level – between these uses more accurately than the functional roles of e.g. *beneficiary* and *recipient*. For instance, although both the possessive dative (p. 53) and the ethic dative (p. 56) relate to the *experiencer* function, the semantic distinction between them is clear-cut. As a whole, the chapter provides insights into a topic every student of Spanish is bound to pay attention to.

Chapter 3, written by Triin Lõbus and Jukka Havu, provides a look at the different verbal periphrases of Spanish. The authors successfully treat lexical and grammatical aspect in connection with the key periphrastic structures of the language. The existence and frequent use of different verbal periphrases is characteristic of Spanish, and may also be a learning challenge, especially in the case of constructions with only slightly different or, in some cases, overlapping uses. This is obvious e.g. with <*ir / venir / andar* + gerund>, often without established construction-type equivalents in other languages. The authors offer pertinent accretions to the theoretical framework laid by García González (1992), for instance when pointing out that <*andar* + gerund> is often used to present the action as unimportant, unnecessary or inappropriate (p. 92).

In Chapter 4, Johan Falk takes a look at *ser* and *estar* ('to be') as copulative verbs. The presence of two *be* verbs in Spanish is a challenge for foreign learners. When the author points out that nouns corresponding to events take *ser* in locative settings (p. 106), it might be worth mentioning – for pedagogical reasons – that this also applies to location in time (e.g. *La reunión es a las dos* 'The meetings is at two o'clock'). The *para reflexionar* box, on page 105, resumes the key differences between Spanish

and other Romance languages regarding the equivalents of *ser* and *estar*. While what the author says about French, Portuguese and Catalan holds absolutely true and is useful background information, one might slightly disagree with him about Italian. Indeed, while *essere* can be used in all cases in Italian as the equivalent of *be*, *stare* has some uses very similar to those of *estar* in Ibero-Romance. According to the dictionary of Collins-Sansoni (1981: 2132–2133, s. v. *stare*), this verb expresses, among other things, location (*le montagne stanno a nord*), health and economic conditions, etc. (*sono stato male tutta la notte*), habitation (*stiamo al secondo piano*), and stay in a specific place (*stare al sole*).

Spanish is rich in passive constructions. In Chapter 5, Susana S. Fernández deals with the characteristics and uses of each of them. However, the reader would also have expected to see a mention of the non-reflexive use of the third-person plural with active predicates in this context (e.g. *Aquí venden coches* ‘They sell cars here’). The author successfully defines the key difference between the periphrastic passive (*ser* + aux.) and the reflexive passive (*se* + verb in 3<sup>rd</sup> p.): the former normally adopts the perspective of the patient while the latter focuses on the event itself. French has a periphrastic passive (*être* + aux.) as the first option (apart from *on* + 3<sup>rd</sup> p. sg.), while the French equivalent of the Spanish reflexive passive is of rather limited use. For this, students also familiar with French may tend to use the periphrastic passive excessively while writing Spanish. Therefore, it might be a good idea to mention this e.g. in one of the *para reflexionar* boxes.

Chapter 6 (Kåre Nilsson, Ingmar Söhrman, Santiago Villalobos and Johan Falk) is dedicated to copulative verbs expressing change. Again, this is potentially a difficult topic for a foreign learner, because unlike languages such as English, French or Swedish, which virtually use one single verb for this purpose (*become*, *devenir*, *bli*), Spanish uses several verbs. Overall, the chapter can be regarded as an excellent introduction to these questions.

Chapter 7, written by Ana Beatriz Chiquito, is about modality. The author focuses on three kinds of modality, epistemic, deontic and dynamic, and discusses how they manifest in Spanish. One interesting detail is that, unlike e.g. García González (1992), the author (pp. 187–189) classifies <*poder* + INF> as a verbal periphrasis, instead of treating it as a “normal” construction, consisting of an auxiliary combined with a verb in infinitive. On the other hand, the Chiquito’s interpretation is shared e.g. by Martínez Gómez (2004) and, recently, by *Nueva gramática de la lengua española*

(2010: 2141). After all, the criteria for a construction to be or not to be considered as a verbal periphrasis are somewhat vague and partially open to interpretations. Otherwise, the chapter is a concise and easily intelligible presentation of the key concepts of modality.

In Chapter 8, Jeroen Vandaele and Mieke Neyens examine the Spanish subjunctive. The chapter, of some 40 pages, provides illustrative examples of variation between the indicative and the subjunctive in cases where both are viable options (pp. 207 and 212). For foreign learners (except for native speakers of a Romance language), the subjunctive mode is a learning challenge because, in many languages, it often does not exist as a separate, morphologically marked category. For this, foreign learners tend to replace it with the indicative. The chapter is recommended reading for any student, because it may help them reach a deeper understanding of this complex topic.

In Chapter 9, Johan Falk discusses the Spanish adverbial gerund. On page 239, he points out the impossibility of using the gerund in Spanish to replace relative clauses (*\*un marino cantando* ‘a singing sailor’). Since many Spanish learners are also familiar with French, it might be a good idea to add here that the above use is, by contrast, completely grammatical in this language (*un marin chantant*). On the other hand, there is one famous exception to the ungrammaticality of the use of the gerund in these cases in Spanish: *el agua hirviendo* (‘boiling water’), which could be mentioned in the chapter as an interesting isolated case. All in all, the chapter is a concise and useful description of the Spanish adverbial gerund.

In Chapter 10, Erla Erlendsdóttir examines the principal simple prepositions of Spanish. This otherwise clear and illustrative presentation could be complemented by a few details. The first one is the use of preposition *de* to express what is called the *essive* case in some languages (e.g. Finnish), mainly corresponding to *som* in the Scandinavian languages (or, occasionally, to some other preposition), e.g. *estar de camarero* (Sw. ‘jobba *som* kypare’, Fi. ‘työskennellä tarjoilijana’), *de postre* (Sw. ‘till efterrätt’, Fi. ‘jälkiruoaksi’). In fact, *de joven* in example (82), p. 275, is used with the above meaning (Sw. ‘*som* ung’, Fi. ‘nuorena’), which is not limited to the age of a person. Secondly, when dealing with the temporal uses of preposition *a* (p. 280) the author could also mention its use in time expressions including adjective *siguiente* (‘following’) (*al día / año siguiente / a la semana siguiente* ‘on the following day’, ‘the following year’, ‘the following week’).

Finally, in Chapter 11, Alejandra Donoso deals with verbs of movement, focusing on differences in strategies between languages, e.g. Spanish and Swedish, in the construction of verbal expressions of movement. The author illustrates, based on the typology of Talmy (2000) how Spanish expresses motion with the main verb and Swedish (like other Germanic languages) with a satellite (e.g. a particle): *entrar – gå in*. By contrast, manner is expressed in Spanish in a gerund or an adjective, prepositional or adverbial clause and in Swedish in a main verb, but also in gerunds and adverbs: *el pájaro salió volando – fågeln flög ut*. The learners of Spanish become implicitly aware of these differences, but it is useful to have a look at them also explicitly at some point. Even though the structural differences in question are considerable, I believe the problems dealt with in the previous chapters are much more challenging for the learner.

After the text chapters, there is an index (pp. 311–317) of the key concepts treated in the book; this significantly helps the reader use the book as a reference.

To conclude, *Temas de gramática española para estudiantes universitarios* is a book that without doubt provides an interesting contribution to studies of Spanish at university. It will also allow for future periodic updating.

## References

- García González, Javier (1992) *Perífrasis verbales* [Verbal periphrases]. Madrid: SGEL.
- Collins-Sansoni (1981) *Italian Dictionary*. Third edition revised and enlarged. Firenze: Sansoni Editore.
- Martínez Gómez, Esther (2004) Las perífrasis verbales del español [The verbal periphrases of Spanish]. *Revista electrónica de estudios filológicos* 7. <<http://www.um.es/tonosdigital/znum7/estudios/kdelasperifrasis.htm>> (4 August 14).
- Nueva Gramática de la Lengua Española* [New grammar of the Spanish language] (2010) Segunda tirada corregida. Madrid: Real Academia Española.
- Talmy, Leonard (2000) *Toward a Cognitive Semantics*. Cambridge: MIT Press.

## Contact information:

Ilpo Kempas

University of Helsinki & Seinäjoki University of Applied Sciences

e-mail: [ilpo\(dot\)kempas\(at\)helsinki\(dot\)fi](mailto:ilpo(dot)kempas(at)helsinki(dot)fi), [ilpo\(dot\)kempas\(at\)seamk\(dot\)fi](mailto:ilpo(dot)kempas(at)seamk(dot)fi)

Harvey, Kevin (2013) *Investigating Adolescent Health Communication. A Corpus Linguistics Approach. Research in Corpus and Discourse*. Bloomsbury: Bloomsbury Publishing Plc. Pp 242.

Reviewed by Pirjo Salomaa

## 1. Introduction

*Investigating Adolescent Health Communication* is, as Professor Ronald Carter puts it in the preface of Harvey's book, a book where empirical real-world problems are demonstrated by describing language and communication in the medical world. Harvey analyzes data collected from emails sent by adolescents. These emails, anonymously sent to the website Homepage of Teenage Health Freak, comprise a collection of over 2 million words.

Harvey's work is the first book-length study in which corpus linguistics is used in interrogating adolescent health discourse and, therefore, has much to offer to linguists and health practitioners who want to improve communication with young people. The book provides valuable information such as teenagers' understanding of sexually transmitted diseases: Condoms are considered a form of birth control rather than protection against sexually transmitted diseases.

The aim of the study presented in Harvey's book is to outline the main discourses of the health concerns of teenagers as online clients to professionals working with young people. The secondary aim is to evaluate how relevant the findings of the study are to professionals and educators working with adolescent health concerns.

The book is divided into 10 chapters. Chapters 3–7 and 9 are mentioned in this analysis because these chapters avail the reader of the opportunity to more closely examine the two main aims of Harvey's study. Harvey uses mixed-methodology when analyzing the data collected from the website. The corpus consists of verbatim questions from young people to online professionals. The questions teenagers pose are related to sexual and mental health, topics which are not easy to discuss in a patient-professional setting. The linguistic material can as a discourse be considered a "window into teenagers' social and mental worlds" because

the material has been left unchanged, including the typographical and spelling errors produced by the teenagers (Cameron 2001: 17).

## **2. The Teenage Health-Freak website as corpus**

The Teenage Health Freak website is popular, with about 8,280,000 hits up until 2012. The language focus is email language and the whole corpus is known as the Adolescent Health Email Corpus (AHEC).

Young people write to the online Primary Care Physician, Dr. Ann, and consult the doctor. The situation is akin to a virtual medical practice, as young people face the same problems as they would in a face-to-face doctor-patient setting. Problems that might arise include young people's concerns that the doctor might not keep the exchange of information confidential. Another communicative obstacle is that the consultations are short (about 2 minutes) and young people may fail to ask or adequately express questions because they lack basic knowledge of the issues at hand. Young people report that professionals might use medical terminology that is difficult to understand and generally concentrate on teenagers' physical conditions rather than treating them as individuals.

The asynchronous communication might help teenagers to feel safe because they are accustomed to sending emails, which allow writers to express intimate thoughts yet remain at a distance. Despite the many advantages of online communication with doctors, limitations persist. Doctors cannot examine their patients physically online, or obtain important paralinguistic information about the patients' health. Nevertheless, emails as a mode of discussing medical issues provide an opportunity for young people to anonymously consult professionals online.

## **3. Adolescent health in context**

Chapter 3 presents adolescence health as a field of research. Adolescence is a period of transition from childhood to adulthood, a time that brings up issues of sexual maturation for discussion. Sexual and psychological problems are sensitive areas to discuss and young people experience difficulties speaking about these issues. Teenagers lack knowledge of reproductive physiology and how diseases are sexually transmitted. Using online health data can offer young people the possibility of gaining

knowledge of psychological and sexual health concerns without editorial interference.

Adolescent mental health is also discussed in the context of suicide, self-harm and depression. According to the World Health Organisation (2000), depression is internationally recognized as a significant contributor to suicide. Psychiatry tends to define mental illness in terms of biological origins. Mental health experiences are rarely discussed, a fact that makes the qualitative research of young people's self-destructive behavior worth investigating. Harvey's analysis of the terminology young people use when they write about their psychological stress provides evidence of how stigmatized and taboo this issue is.

#### **4. Methods and Data**

In Chapter 4, Harvey introduces his corpus-assisted analysis, which is both quantitative and qualitative with insights from medical sociology, psychology and the health sciences. The corpus is, in other words, a collection of machine readable and authentic texts combining terminology of medical sociology, psychology and health sciences (McEnery & Wilson 2001: 5). Health care has attracted attention from health services, ethics, psychology, social sciences, anthropology, media studies and linguistics.

The emails anonymously sent to specialists Dr. Ann McPherson and Dr. Aidan Macfarlane between 2004 and 2009 created a corpus of 113 480 messages, or 2 million words.

#### **5. Frequency and keyword information of Adolescent Health Email Corpus (AHEC)**

Chapter 5 presents the frequency and keyword information of the AHEC. In total, teenagers send 400 requests for health information to the Teenage Health Freak website every week. By using frequency and keyword techniques some significant characteristics were revealed. Personal pronouns were used often, reflecting the self-oriented nature of emails. Third-person terms she and he (as well as friend and mate) were frequent, probably because young people were willing to use third-person references as a stylistic form in place of a first-person pronoun. In his analysis Harvey used keyword groupings and made some connections within semantic domains such as sexual health. For example these keywords referred to

sexual health: *sex, pregnant, period, condom, pill* etc. Words that referred to sexually transmitted infections were for example: *STI, AIDS, chlamydia, STD* and so on. Keyword patterns around the theme mental health were terms like *depression, depressed, cutting, harming, suicide, suicidal*, expressing the disorder and self-injurious behavior. The health-associated keywords do not cover all the issues related to sexual health and mental health, but they are lexical elements that are worth closer examination.

## **6. Qualitative approach to the analysis of the data: Reproductive health concerns**

In Chapter 6 a selective analysis is used to identify sexual health related themes from collocates. The table on page 101 presents left-hand and right-hand collocates for the keyword *sex*, indicating that the most regularly collocating words are functional words such as *have, having, you is, can* etc. The lexical collocates provide a picture of various contexts in which they are used: *want, unprotected, condom, pregnant* etc. The analysis of data also highlights that the discourse illustrating sex can be considered “male-sex drive discourse”, where female teenagers are “objects”. This is evident in the occurrence of emails written by young women where the expression *my boyfriend wants* is frequently used. At the same time, the discourse concerning the issue of sexual activity of the male might refer to heterosexual penetrative intercourse. Further, analysis of reproductive health, folk beliefs and misinformation was expressed in the data. Collocational analysis of the terms *condom* and *condoms* revealed that young people considered condoms a form of birth control rather than a form of protection against sexually transmitted diseases.

## **7. Questions about sexual transmitted infections: HIV/AIDS**

Chapter 7 discusses the stigma and discourse of HIV and AIDS. In the data collected from young people’s emails, HIV/AIDS are the most commonly mentioned sexually transmitted diseases. *HIV/AIDS* are associated with *uncleanliness, pollution* and *death*. Further, a miasmatic discourse still persists, and the terms HIV and AIDS still produce confusion among teenagers. Teenagers were not aware of how HIV/AIDS could be contracted: Coming into contact with infected objects and even acquiring the virus from someone who did not have it were frequently expressed

beliefs. Further, young people used the verb ‘to catch’ HIV/AIDS which is used to encode contracting common infections. The other verb choice, ‘to get’ HIV/AIDS, also expresses that HIV/AIDS is a highly infectious disease. In other words, the understanding of the fact that HIV/AIDS is transmittable only through specific activities has not reached every teenager.

## **8. Communication of psychological distress: Suicide and self-harm**

In chapter 8 the author presents keywords under the rubric of mental health and disorder. The ten most frequent words are: *depressed, die, harm, depression, sad, suicide, cutting, wrists, unhappy, suicidal*. The number of keyword expressions relating to suicide is 12, 6 of the most often used expressions being: *suicide, kill myself, want to die, overdose, suicidal and end it*. All these words and expressions can be considered (strong) expressions of mental disorder caused by different reasons in young people’s lives. As a result of Harvey’s corpus research concerning the communication of psychological distress, reasons for suicidal impulses expressed in adolescents’ emails are family dysfunction, bullying, eating disorders and body image, alcohol and drugs, stress, sexuality and anxiety. Teenagers seem to be unable to seek help or resolve their own problems. Cutting and self-harming may be a way for teenagers to gain emotional relief in problematic situations. Self-harming also tends to become a habit, an addiction hard to get rid of without professional help. Further, professionals do not necessarily have a sufficient understanding that young people cannot rationalize their mental disorders and are in need of help.

## **9. Adolescent accounts of depression**

Chapter 9 discusses defining *depression* and the expression *depressed*. Researchers define depression as a descriptive noun, a vague term for a variety of states (Wittink, Dahlberg, Biruk & Barg 2008: 171). Clinically, there are diagnostic criteria used in diagnostic classification of depression: DSM-5, 2013 (Diagnostic and Statistical Manual of Mental Disorders) and ICD-10 (International Statistical Classification of Diseases and Related Health Problems 10<sup>th</sup> revision) used by professionals. These clinically used methods for diagnosing depression based on symptoms experienced by

people are not easy to analyze, because the symptoms of depression vary periodically and individually, even daily.

Linguistically, there is an interesting difference between the expressions: I *am* depressed and I *have* depression. In English the first of these expressions might describe a state caused by, for example, bullying or a reaction to the personal and social contexts of (young) people. The latter expression refers to an organic and pathological state, or a state where deeply ingrained psychological problems exist. The linguistic analysis of young people's expressions concerning depression in their emails gives a picture of chaotic experiences hard to objectify and express.

## 10. Conclusions

Investigating adolescent health communication aims to give a demonstration of a corpus approach to adolescent health communication as a discourse and succeeds in this aim. Harvey's method is a combination of quantitative and qualitative approaches where keywords, collocations, and concordances are used in order to outline the findings in the study. The book is well-structured, containing tables and examples from emails sent by young people. The other aim of this book is, as Harvey expresses, to make data collected in this research useful in communication training programs. This aim is a greater challenge because of the limited amount of readers who may have experience with adolescent health concerns as professionals, and be familiar with a corpus linguistics approach. For professionals in the know, this book gives a larger perspective and new ideas suitable for improving communication with adolescents in different contexts. Professionals in health education could elaborate upon Internet programs in which they could use the language teenagers use while discussing sexual health and psychological stress. Using pictures and concrete medical and scientific terminology would assist professionals in reaching a larger group of teenagers and avail patients of the possibility of informative, dialogic and "good-quality" online communication.

## References

- Cameron, Deborah (2001) *Working with Spoken Discourse*. London: Sage.  
*Diagnostic and Statistical Manual of Mental Disorders* (DCM-5).  
<<http://www.psychiatry.org/dsm5>> (21 November 2014).

- International Statistical Classification of Diseases (ICD-10)*. <<http://www.who.int/classifications/icd/en/bluebook.pdf>> (21 November 2014).
- McEnery, Tony & Wilson, Andrew (2001) *Corpus Linguistics*. Edinburg: Edinburgh University Press.
- Wittink, Marsha; Dahlberg, Britt; Biruk, Crystal & Barg, Frances (2008) How older adults combine medical and experimental notions of depression. *Qualitative Health Research* 18: 1174–83.
- World Health Organisation. (2000) *Mental Health and Brain Disorders*. Geneva: World Health Organisation.

#### Contact information

Pirjo Salomaa  
Centre for Applied Language Studies  
P.O. Box 35  
40014 University of Jyväskylä  
Finland  
e-mail: pirjo(dot)i(dot)salomaa(at)student(dot)jyu(dot)fi

## Contents:

- 5 External Reviewers of *SKY JoL* 27 (2014)
- 7 Thomas Groß  
Some Observations on the Hebrew Desiderative Construction – A Dependency-Based Account in Terms of *Catenae*
- 43 Wojciech Lewandowski  
Deictic Verbs: Typology, Thinking for Speaking and SLA
- 67 Ulla Vanhatalo, Heli Tissari & Anna Idström  
Revisiting the Universality of Natural Semantic Metalanguage: A View through Finnish
- 95 Mari Wiklund  
La transmission des effets stylistiques des phrases sans verbe fini dans les traductions finnoises – L'exemple des pièces de théâtre de Jean-Paul Sartre
- 127 Peter Wikström  
#srynotfunny: Communicative Functions of Hashtags on Twitter

### Squibs:

- 153 María Luisa Carrió-Pastor  
Cross-cultural Variation in the Use of Modal Verbs in Academic English

### Book reviews:

- 167 Maurer, Philippe (2009) *Principense (Lung'le)*. Grammar, Texts, and Vocabulary of the Afro-Portuguese Creole of the Island of Príncipe, Gulf of Guinea
- 173 Susana S. Fernández & Johan Falk (eds.) (2014) *Temas de gramática española para estudiantes universitarios, Una aproximación cognitiva y funcional*
- 179 Harvey, Kevin (2013) *Investigating Adolescent Health Communication. A Corpus Linguistics Approach. Research in Corpus and Discourse.*

### Publisher:

Suomen kielitieteellinen yhdistys  
Språkvetenskapliga Föreningen i Finland  
The Linguistic Association of Finland