



---

## VIENNA ENGLISH WORKING PAPERS

---

VOLUME 18

NUMBER 2

DECEMBER, 2009

INTERNET EDITION AVAILABLE AT:  
[HTTP://ANGLISTIK.UNIVIE.AC.AT/VIEWS/](http://anglistik.univie.ac.at/views/)

---

### CONTENTS

LETTER FROM THE EDITORS .....	1
<b>OLGA FISCHER</b>	
Grammaticalization as analogically driven change? .....	3
<b>KAMIL KAŻMIERSKI</b>	
On the role of perception in the acquisition of the <i>peach</i> – <i>pitch</i> contrast by Polish learners of English .....	24
<b>EVA MARIA SEREGÉLY</b>	
“I know this word. (I think) It means ...” – Lexical competence and self-assessment in CLIL and traditional EFL classrooms .....	42
Impressum .....	70

---

### LETTER FROM THE EDITORS

*Dear Readers,*

As we enter what shall hopefully turn out to be another decade full of interesting and inspiring *Views* on a wide variety of subjects, it is time for a brief reflection on things past. Over the expiring decade, *Views* has brought you a total of 22 issues, containing a staggering 150+ articles (including all special issues) by 170 different (co)authors. Quite an achievement, even if we say so ourselves! To round off the decade, we now add the present issue, which once more underlines the breadth of our journal's coverage, showcasing

contributions on the topics of language change, speech perception, and language learner (self-)assessment.

In the first article, Olga Fischer, one of our long-time outstanding international contributors, returns with a thought-provoking, critical piece on recent theoretical conceptualizations of 'grammaticalization' as a process of language change. Advocating a grounded perspective that takes the speaker-listener's linguistic and cognitive reality into account, she argues that grammaticalization constitutes a heuristic of diachronic analysis rather than an intrinsic, unified, unidirectional process guiding linguistic change; and she stresses and expounds on the centrality of the role of analogy in this regard.

The second article introduces first-time *Views* contributor Kamil Kaźmierski to our readership, who presents his research on the role speech perception may play in certain English language acquisition and production difficulties experienced by native Polish speakers. Based on an empirical study that includes a perception experiment, Kaźmierski tests, but ultimately finds reason to reject, the hypothesis that the English vowels /i:/ and /ɪ/ are perceptually assimilated by Polish speakers due to features of their own native language, which was suggested to account for previously attested problems in the rendition of the contrast.

Eva Maria Seregély, who is also featured in our Special Issue "Bridging the Gap between Theory and Practice in English Language Teaching" (*Views* 18/3), completes the present issue, now providing an in-depth report on her comparative research assessing the English competence of CLIL ('Content and Language Integrated Learning') vs. 'traditional', non-CLIL students in an Austrian grammar school. Focusing on results from a self-report test on lexical competence administered to both groups, Seregély finds indication for the fact that CLIL students indeed benefit from the special program, as they outperform their peers from regular programs in this respect.

As always, we hope that you will enjoy the papers presented in this *Views* winter issue and find them stimulating and inspiring. As always, we would like to hear from you with any comments you may have.

Here's to a successful 2010, and to another decade of lively scholarly exchange, discussion, interaction, and insights, through sharing our *Views*!

## THE EDITORS

# *Grammaticalization as analogically driven change?*

*Olga Fischer, Amsterdam\**

One of the ubiquitous principles of the psycholinguistic system is its sensitivity to similarity. It can be found in the domains of perception and production (as well as learning).

The more similar any two sets are, the more likely the wrong rule is applied.

(Berg 1998: 185; 236)

## 1. Introduction

Since the 1980s, grammaticalization has been a popular research topic in diachronic linguistics, with its field of application widening considerably over time so that the phenomenon of grammaticalization came to be elevated to theoretical status: a model to understand how language is used and structured, and develops through time. Its spreading popularity has also led to increasing concern about quite a number of aspects related to the model. Some of the more important questions raised are:

- (i) Is grammaticalization an independent mechanism or an epiphenomenon (i.e. a conglomerate of changes occurring elsewhere that happen to coincide in cases of grammaticalization)?
- (ii) What is the relation between the synchronic speaker-listener and the essentially diachronic nature of grammaticalization? What role is played by the synchronic system that the grammaticalizing structure is part of?
- (iii) What empirical evidence do we have for grammaticalization, and, perhaps more importantly, where should we look for evidence?
- (iv) What causes grammaticalization and language change in general? Should the mechanisms that apply in language learning also apply in language change? And more particularly, what is the role of analogy, reanalysis, frequency, to mention some of the more important factors, in this context?<sup>1</sup>

---

\* The author's email for correspondence: [O.C.M.Fischer@uva.nl](mailto:O.C.M.Fischer@uva.nl).

<sup>1</sup> Language contact is also relevant here, but since this is an external rather than an internal factor it will be largely ignored.

These questions will be addressed in section 2, where I will also highlight the position of analogy. In section 3, I will explore the nature of analogy further in order to find out in how far grammaticalization can be understood as an instance of analogically driven change.

## 2. The nature of grammaticalization

In order to know what grammaticalization is we have to determine what its characteristics are. Problematic here is that the process covers quite a number of phenomena, and different ones for different linguists. Its core for most linguists, roughly following Meillet's (1912: 131) early definition, involves the development of a particular lexeme (or a combination of lexemes as part of a construction) into a grammatical function word, on a 'cline' from lexical to grammatical, where the development could also pertain to any subpart of this cline. Gradually, however, grammaticalization began to include the development of grammatical constructions in general, without the kernel of substantive elements, so that general word-order restrictions or the creation of new syntactic patterns also became part of it; witness e.g. Givón's (1979) statement that syntax develops out of discourse, as well as his treatment of clause-combining (see also Hopper/Traugott 2003: 194; 209-11; Bybee 2003). Other clines having to do with subjectification (where the cline moves from propositional via textual to more epistemic stances) were likewise seen to fall under grammaticalization (cf. Traugott 1982, 1989, 1995).

This widening led to a weakening of the power of grammaticalization as a clearly circumscribed mechanism in change. This can easily be seen from the fact that the parameters originally set up by Lehmann (1982: 306) to characterize the canonical type no longer all neatly have to apply in each case.<sup>2</sup> Lehmann's parameters give the process a unity in that they all involve reduction or loss on both the paradigmatic and the syntagmatic plane, i.e. loss of weight (phonetic attrition, semantic bleaching and scope decrease), loss of choice (paradigms of possibilities become reduced and elements become bonded together), and loss of freedom (elements become obligatory in the clause and fixed in position). Obviously, the development of fixed word order or new syntactic patterns doesn't involve phonetic attrition (unless one thinks

---

<sup>2</sup> It is therefore perhaps not surprising that Lehmann (2004: 155), firmly sticking to his parameters, does not consider the "creat[ion of] new grammatical structure" an instance of grammaticalization. Bybee (2003: 146), although taking a much wider view of grammaticalization (it includes "the creation of word order patterns"), interestingly enough, virtually follows Lehmann's parameters when she discusses the cognitive processes involved in grammaticalization.

of this as whole elements being elided; but note that this would disrupt the widely accepted notion of grammaticalization being gradual) or bleaching. Furthermore, it has been suggested that in many cases, especially those involving subjectification, there is scope *increase* rather than *decrease*.<sup>3</sup> Similarly, in the case of clause-fusion or syntacticization discussed by Givón (1979: 214) – he suggests that complement clauses with non-finite verb forms and PRO subjects may have developed from paratactic clauses with finite verbs and lexical subjects – there is no question of Lehmann’s parameter ‘increase in paradigmaticity’ applying since such constructions usually remain in use side by side.

The widening of the field of application and the consequent weakening of the parameters would not be a matter of concern if grammaticalization is considered merely a heuristic device, enabling us to spot the process at work – with only half the parameters being present one could still discover this. However, for many grammaticalizationists it is more than that, it is a unified, unidirectional development that guides, and hence explains, change; it cannot be cut up into pieces, and indeed unidirectionality was and still is one of its main principles (cf. e.g. Bybee 2003: 145; Lehmann 2004: 154; Kiparsky forthcoming). For these linguists, cases of degrammaticalization were thought not to exist or were explained away because they did not square with the notion of unidirectionality, while the (opposite) process of lexicalization came to be seen as a different, orthogonal phenomenon even though the processes working in both grammaticalization and lexicalization are in fact very similar (cf. Himmelmann 2004, Fischer 2007 pace Lehmann 2004: 168-70, Brinton/Traugott 2005).<sup>4</sup>

Linguists, with formal, functional, as well as more philological backgrounds, who combined their voices in the critical volume of *Language Sciences* 23 (cf. also Janda/Joseph 2003, Joseph 2004), stress the fact that all the changes occurring in grammaticalization may also occur independently, thereby querying the nature of the unity of Lehmann’s parameters from

---

<sup>3</sup> Tabor/Traugott (1998) and Roberts/Roussou (2003) even see scope *increase* as the rule in grammaticalization, referring to the development of modals from full verbs into auxiliaries and from deontic into epistemic modals, and the rise of pragmatic markers (for counterarguments, see Fischer 2007).

Increase should be the rule in cases of degrammaticalization, as can indeed be seen in the development of the infinitival marker *to* in English, in contrast to German *zu* and Dutch *te* where scope decrease has occurred (cf. Fischer 1997 pace Haspelmath 1989 and Bybee 2003, who consider this case a normal case of grammaticalization in all three languages).

<sup>4</sup> This difference is constituted not by the process but by the lexical source input. Only very general items of a basic nature grammaticalize (cf. Bybee 2003: 151).

another direction. These linguists generally stress that more attention should be paid to the speaker-listener, and the synchronic language system used to produce or interpret language utterances. This is not to say that in grammaticalization theory no attention is paid to the speaker-listener level but this is mainly confined to the immediate pragmatic-semantic context, while the shape of the (formal) system also guiding the speaker-listener is ignored (cf. Mithun 1991, Fischer 2007). In general, supporters of grammaticalization see the process as being driven by pragmatic-semantic forces only, a “product of conceptual manipulation” with changes in form *resulting* from this (Heine et al. 1991: 150;174; and cf. Hopper 1991: 19; Rubba 1994; Hopper/Traugott 2003: 75-6). In other words, they would not admit the possibility of form also driving a change.

Looking at grammaticalization from a purely synchronic, speaker-listener point of view rather than a diachronic one may shed a different light on the process or mechanism called ‘grammaticalization’. Even though diachrony is present in synchrony in the form of variation, it is not the case that a ‘pure’ synchronic system does not exist, as Lehmann (2004: 153) maintains. For the speaker-listener, there *is* only the synchronic system at any moment of speech.<sup>5</sup> The point is that the speaker-listener has no panchronic sense, he doesn’t necessarily see the connections between the grammaticalization variants in a *historical* light.<sup>6</sup> In other words, in order to prove the existence of grammaticalization as an *actual* mechanism of change linked to human processing, one cannot fall back on the historical process itself. However, this is what is typically done in grammaticalization studies. These attempt to empirically prove the ‘reality’ of grammaticalization as a mechanism by showing its universal pathway (cf. Haspelmath 1989, 1998,<sup>7</sup> Heine 1994,

---

<sup>5</sup> This does not mean that the synchronic system cannot change in a speaker’s lifetime; it does constantly, i.e. change is not confined to the period of acquisition only. The point is that in actual speech situations the speaker uses a stable synchronic system.

<sup>6</sup> It is interesting to note in this connection, that even generative linguists have been guilty of panchronism, cf. the synchronic phonological ‘rules’ suggested by Chomsky/Halle (1968) to derive the stem-vowels in such items as divine/divinity from the same underlying phoneme [i:], neatly following the actual historical development of these vowels from earlier [i:]. From the processing point of view, it seems rather more likely that such words are learned holistically rather than rule-based. Recent studies on the workings of the mind/brain emphasize that retrieval from memory is the preferred strategy, and that people are able to store vast numbers of prefabricated units (cf. Pulvermüller 2002: 193, Dąbrowska 2004: 27).

<sup>7</sup> Haspelmath (1999) still considers grammaticalization unidirectional. However, he no longer sees it as an independent mechanism but a side effect of the maxim of extravagance.

Bybee 2003): the “diachronic identity” or “continuity of two forms or constructions  $F_1$  and  $F_2$ , at  $T(\text{ime})_1$  and  $T_2$ ” (Lehmann 2004: 156ff.).<sup>8</sup>

Now this may constitute empirical evidence if one looks at change on the *language output* level: the diachronic stages may be seen as connected, with the constructions at each stage changing gradually, almost imperceptibly, by pragmatic inferencing, analogical extension and reanalysis. However, this scenario need not have any reality at the *processing* level, where the same constructions need not be connected at all. The following question should be raised: is there an actual reanalysis in psycho-/neurolinguistic terms? This point is important considering the fact that it is ultimately the speaker-listener who causes the change.<sup>9</sup>

The ‘grammaticalization’ of constructions, or the way (diachronically connected) forms are stored in our brains could be said to resemble the process of conversion, and their storage. When a noun like *table* is used as a verb, the two items are stored in different paradigms or categories, both formally and semantically, and, once there, they may drift further apart. There is no question of reanalysis here for the speaker-listener; he is simply making use of the (abstract) grammar system of English that allows such an option (and with increasingly greater ease after most inflexions were lost in the Middle English period). Since there are many such hybrid items in the language, he analogizes, on the basis of an existing pattern, that *table* belongs to this pattern too. How is he to know that *table* had not been used as a verb before, when this verb=noun scheme is such a common pattern in his language?

In a similar way, with the construction *going-to*+infinitive, a present-day speaker-listener identifies it in any actual speech situation as either a full lexical verb followed by a purposive *to*-infinitive, or an auxiliary (with *to* incorporated) followed by a bare infinitive, according to the patterns of the full verb and the auxiliary paradigms that he has mastered in the course of language acquisition. As with conversion, the speaker-listener doesn’t reanalyse, he categorizes holistically, whereby he may apply the ‘wrong’ rule.

---

<sup>8</sup> Berg (1998: 51) writes: “reference to a historical development is insufficient as an explanation of a given fact. It should always be asked why a certain development follows one path rather than another and why it took place at all.”

<sup>9</sup> Analogy works in the same way as the perception of prototypes, i.e. only a certain amount of similarity is necessary for an item to be seen as belonging to a particular (structural) type. In this sense a structure or word may “snap into place” as if it were a prototypical member of an existing category (Denison 2006: 281). My hypothesis, pace Denison, is that “snapping into place” is more likely to occur in processing “than gradience”: the creation of a new (intermediate/gradient) constructional category.

How he categorizes in each case depends on the present state of his grammar as well as the context, just as he can recognize whether *table* is a noun or a verb from the (syntagmatic) context and the paradigmatic inventory of patterns present in his grammar. The context is characterized by formal (i.e. position, word order, the presence of a determiner, inflections etc.) as well as semantic-pragmatic information. The very first time a historical speaker-listener identified *going-to* as auxiliary, therefore, did not constitute an actual reanalysis of *going*(full verb)+*to*-infinitive but a category mistake, a mistake that he could make because the *going-to* form fitted both the V-*to*-V as well as the Aux-V pattern.<sup>10</sup>

Analogical extension is similar, too, in terms of speaker-listener processing: like grammaticalization and conversion it is also based on pattern recognition and categorization. When a speaker uses *brung* rather than *brought*, or *shaked* rather than *shook*, there is no question of reanalysis. He uses past-tense *brung*, because it fits another past-tense pattern: *rung*, *stung* etc., which happens to be far more frequent than the pattern of *brought*.<sup>11</sup> The important point about analogical extension is that it occurs proportionally. It doesn't simply involve the "expansion of contexts in which a construction can occur", "adding new peripheral members [e.g. new infinitives, inanimate subjects] to a category [e.g. *going-to*]" (Bybee 2003: 158); it happens because, once *going-to* is interpreted as belonging to the Aux-category, it will follow the behaviour of other members of this new category.

In all three cases, we can thus provide a historical explanation for the new forms. However, although a certain overall continuity or development

---

<sup>10</sup> The idea of a 'mistake' is an important point that Deacon (1997: 74) makes with respect to the recognition of iconicity (or analogy):

*Usually, people explain icons in terms of some respect or other in which two things are alike. But the resemblance doesn't produce the iconicity. Only after we recognize an iconic relationship can we say exactly what we saw in common, and sometimes not even then. The interpretive step that establishes an iconic relationship is essentially prior to this, and it is something negative, something that we don't do. It is so to speak, the act of not making a distinction.*

The analogy between *going to* and the Aux-paradigm is due to partial (but enough) similarity in meaning (sense of future/possibility) as well as position (adjacent to the infinitive). Note that the formal similarity involves an abstract pattern, not a similarity in lexical or phonetic (surface) form. It is interesting to note that *gonna* also follows the Aux-paradigm in that only adverbs like *soon*, *always*, *often* can be placed between *gonna* and the infinitive, which is not possible after fully lexical *going to*.

<sup>11</sup> There are other past forms with the same phonetic shape as *brought*, i.e. *fought*, *caught*, *sought*, but pattern recognition (and hence a categorization 'mistake') here is not to be expected because it is not supported, as in the case of the *rung*-group, by homogeneous shapes in the present tense, i.e. all the *rung*-members have present tense *-ing* forms.

(unidirectionality) may be ascertained – especially with surface forms connected by ‘grammaticalization’, and, on a more abstract level, with strong verbs becoming weak (rather than vice versa<sup>12</sup>) – such unidirectionality need not be the case, as we can see in the case of *brung*. In terms of synchronic processing, the choice is not guided by any historical development but by the strength of the patterns that the form can be seen to belong to, and this strength depends in turn on the frequency of the patterns themselves. If one of the variant forms is more of a grammatical function word (as with *going-to*) or a more basic vocabulary item, then that variant will be more frequent, and may become the norm, often followed by the loss of the older form if there is not enough distinction in meaning to preserve both. It could also be said that this type of processing is in fact no different from our ability to fill a sentence pattern like SVO with different lexical elements chosen from the NP and VP categories. That too is a choice, not a reanalysis each time of the SVO pattern.

If we follow this line of argument and try to understand what grammaticalization entails from the synchronic speaker-listener aspect, then it is not necessarily the case that the ‘cline’ (which has reality only on the level of the historical development of language-output data) has to continue inexorably in the same unidirectional way. Quite possibly, it may, and it often does (due to the fact that the more grammatical variants also happen to become more frequent over time), but it does not always, as shown by attested cases of degrammaticalization, or in cases where weak verbs become strong. Sometimes also processes stop halfway, and similar processes with the same starting point may develop differently in different languages as has happened, for instance, with the modals and the infinitival marker (cf. footnote 3) in Germanic languages.

What may stop a process or what may cause degrammaticalization? It could be a drop in frequency of the item or construction concerned, for whatever reason. But in cases of degrammaticalization, it may be a change elsewhere in the system, which affects the pattern that the grammatical element belongs to. If indeed an important driving force in the grammaticalization of a particular construction is the availability of a grammatical category or pattern that it could fit into, then in a similar way, but with the opposite effect, the non-availability of a pattern may drive

---

<sup>12</sup> In this case there is no item-based continuity as with *going-to* since there can be no gradual change from *brought* to *brung*, only an instantaneous one, the reason being that the choices for *going-to* arose syntagmatically, while those for *brung/brought* arose paradigmatically. In both cases, however, the possibility for the choice lies in the analogy, mistaking one pattern for another. Conversion is like *brung/brought* in this respect, but here there is no unidirectionality at all.

*degrammaticalization*. Plank (1995) has shown that, in the case of the English genitive inflection becoming a clitic, this follows from the fact that the inflectional system of nouns had been eroded so that the genitive ending had become isolated, no longer fitting the new, inflectionless, noun-pattern. A similar situation existed in the case of the Irish 1st plural verb ending *-mid*, which had become the only inflected pro-form in the plural. The fact that *-mid* was upgraded to an independent pronoun, *muid*, is not surprising considering that the pronoun pattern was available in the rest of the verbal paradigm (cf. Kiparsky forthcoming: 28). In such cases, as Plank makes clear, there is a ‘Systemstörung’, which asks for drastic methods on the part of the speaker-listener to keep the language system manageable.

The hypothesis then is that in both grammaticalization and *degrammaticalization* (and in conversion too) the driving force, next to (syntagmatic) context and frequency, is the availability of a fitting – in terms of formal and semantic similarity – (paradigmatic) category or pattern for the new variant in the synchronic system of the speaker-listener. If this is correct, analogical thinking plays a role in all the above cases. Analogy only happens on the basis of an exemplar, which may be a concrete lexical form or a more abstract morphosyntactic pattern.

Kiparsky (forthcoming: 6) agrees that both *degrammaticalization* and grammaticalization are forms of analogical change, which he calls ‘grammar optimization’. At the same time, however, he makes a distinction between the two: *degrammaticalization* is based on exemplar-based analogy, while grammaticalization is a different matter because it is non-exemplar-based. The analogy in the latter case follows “constraints, patterns and categories ... provided by UG” (ibid: 6), and only arises “under a reduced input” (ibid: 11). In this way, Kiparsky can preserve Meillet’s idea that only grammaticalization can create *new* categories, and he can also save the principle of unidirectionality because *degrammaticalization* is now seen as different in nature and is therefore no longer the opposite of grammaticalization.

There are a number of problems with Kiparsky’s proposal. First of all, it is almost too clever: grammaticalization and *degrammaticalization* are said to be the same because they are based on analogy, but are different as far as unidirectionality is concerned.<sup>13</sup> Secondly, it relies on the idea of an innate grammar – of which we do not know the contours – so that the notion of non-

---

<sup>13</sup> Kiparsky (forthcoming: 6) stresses this problem himself when he writes: “From the traditional point of view, the idea of non-exemplar-based analogy is a contradiction in terms: analogy by definition has a model, a pre-existing pattern of the language which is generalized to new instances”.

exemplar-based analogy is not falsifiable, and indeed not explanatory outside its own linguistic model (cf. Fischer 2007: 67-74, and references there). Thirdly, the idea of non-exemplar-based analogy creating new categories is difficult to distinguish from reanalysis, which is seen by many as primary in grammaticalization (cf. Harris/Campbell 1995; Hopper/Traugott 2003: 39, 63-9; Roberts/Roussou 2003). Since Kiparsky's facilitator for grammaticalization is not based on an existing pattern, but on an innate one, it would have to be called 'reanalysis' by anyone whose model doesn't include UG. Thus, Kiparsky is only able to downgrade or "go beyond" (ibid: 19) reanalysis, by putting up empirically invisible UG patterns to base his analogy on. He rejects reanalysis because it doesn't provide an explanation: "labelling a change as a reanalysis, innovative or otherwise, doesn't get at its nature or motivation. For now the claim that grammaticalization is reanalysis remains virtually a tautology" (ibid: 19). In other words, he does not reject reanalysis because it has no reality from the point of view of speaker-listener processing, as I have done above. Fourthly, we end up with two types of analogy, even though ultimately they are both said to fall under grammar optimization (ibid: 6). This, however, is also a problem because it is well-known that exemplar-based analogy is often very local (cf. McMahon 1994: 70-6); such local cases cannot be said to lead to the same form of grammar optimization as the optimization driven by the much more global rules and constraints of UG. It would, therefore, be simpler if it could be shown that analogy works in the same way in all cases.

The positive aspect of Kiparsky's proposal is that it rejects the process of grammaticalization *itself* as a cause or mechanism for change. He emphasizes that the definitions of grammaticalization given in the literature do not work because the different aspects of grammaticalization "do not have to march in lockstep", and because one aspect is not "a necessary consequence" of another; rather, grammaticalization as described "pick[s] out separate and more or less loosely parallel trajectories of change" (ibid: 4).

Below, analogy will be looked at in more detail; attention will be paid to analogy as a deep-seated cognitive principle that is not only relevant to language processing and language change, but also to learning processes outside language. I will stress that analogy is used to categorize, and that categorization involves both concrete and abstract linguistic signs. In addition, the ability to analogize is evolutionary old and present in other mammals too. Finally, it is an important mechanism in language acquisition (cf. Slobin 1985; Tomasello 2003a), and in the processing of language in general (cf. Berg 1998). If we accept that the system of grammar that each of us acquires in life should be an empirical psychological/biological model, and

not some abstract linguistic model that has no relation to our psychobiological make-up, then this system should reflect human processing, and the key to this should be found with the help of advances in neuro- and psycholinguistics. Berg (1998: 278) writes: “The structure of the language is shaped by the properties of the mechanism which puts it to use”. The more the same mechanisms are seen to operate elsewhere, the more persuasive they become.

### 3. Analogy: its nature and the role it plays in linguistic modeling and change

Analogies can be very concrete or quite abstract; that is, an analogy may be based on concrete lexical items as well as more abstract schemas. Analogy is also a highly fluid concept and therefore works quite differently from the type of global rules favoured by generative linguists. Hofstadter (1995: 198) gives an example of the fluidity of analogical thinking on the very concrete level of language use. He describes analogy as ‘conceptual slippage’ and argues that this slippage is important in order to keep language workable and flexible. It is to be preferred to a rigid system:

*And one last example from this genre, perhaps my favorite ... A grocery-store checkout clerk asked me, ‘Plastic bag all right?’, to which I replied, ‘Prefer a wood one ... uhh, a ... a paper one, please.’ Contributing towards this slip might have been the following factors: paper is made from wood pulp, grocery bags are brownish, somewhat like wood and unlike standard paper, they are also considerably ‘woodier’ in texture than ordinary paper is, and plastic and wood are both common materials out of which many household items are made, whereas paper is not.*

*Substitution errors like these reveal aspects of the subterranean landscape – the hidden network of overlapping, blurred together concepts. They show us that under many circumstances, we confuse one concept with another, and this helps give a picture of what is going on when we make an analogy between different situations. The same properties of our conceptual networks as [sic] are responsible for our proneness to these conceptual-halo slips make us willing to tolerate or ‘forgive’ a certain degree of conceptual mismatch between situations, depending on the context; we are congenitally constructed to do so – it is good for us, evolutionary speaking. (Hofstadter 1995: 198)*

As will be seen below, this ‘conceptual mismatch’ also takes place on a more abstract level, that of the system, once patterns have been formed.

Analogical rules are typically not across the board but work in local areas. Analogical learning starts with concrete situations and is based on experience, both linguistic and situational, just like the kind of analogical reasoning that we saw in Hofstadter’s example, which also depends on a situation and on

previous experience. In learning, the analogies may become more and more abstract by means of what Slobin (1985) has called ‘bootstrapping’. That means that abstract patterns deduced from concrete tokens begin to form a system provided these tokens occur frequently enough. The most frequent concrete and abstract patterns (i.e. idiomatic phrases, such as *He kicked the bucket*, and grammatical schemas, such as the English NP consisting of [(Det) (Adj) Noun]) become automatized and will become part of our lexical and grammatical knowledge.

The advantage of a usage-based grammar (i.e. a grammar that is the result of actual learning), such as the one indicated here, is that no distinction is made between lexical items/phrases, and grammatical words/schemas (as in Construction Grammar).<sup>14</sup> Lexical items are learned first; patterns, both concrete and abstract, follow from that. The learning itself takes place by what Slobin (1985) and Peters (1985) have called ‘operating principles’. These are general strategies, based on analogy, on recognizing what is same and what is not-same, and drawing conclusions from that. These same/different operations are performed on linguistic utterances in context, on the form as well as the situated meaning of the utterance, in which frequency plays an important role. The same analogical procedures also provide us with the ability to build up categories (like Noun, Verb) and syntactic structures (cf. Itkonen 2005; Wanner 2006).

In analogy both iconic and indexical forces are important (as is clear, for instance, from the quote from Hofstadter above, when he used “wood” instead of “paper” because paper is made of wood (indexical) and because the bag looks in colour and texture a bit like wood (iconic). The strong interconnections between the indexical and the iconic are clearly indicated in Anttila’s (2003) ‘analogical grid’, whose paradigmatic and syntagmatic axes represent the ‘woof and warp of cognition’. Anttila emphasizes that all linguistic signs (which include both concrete lexical items and structural patterns) are double-edged, they are combinations of form and meaning.<sup>15</sup> Even more importantly, in view of the force of analogy, he stresses that

---

<sup>14</sup> Cf. Tomasello (2003b: 9), who notes in Langacker’s words that language is a “structured inventory of symbolic units” each with their own form and function. He points out that there is no tidy distinction between lexicon and grammar, and that, for instance, idiomatic constructions of the type *Him be a doctor!*, do not fit the lexicon because the construction is productive, nor the grammar of English since it doesn’t follow the rules.

<sup>15</sup> This is also the accepted idea in construction grammar, where constructions from the lowest to the highest levels form a network of intersecting connections (cf. Goldberg 2006: 18; Noël 2008). Similar ideas about the organization and storage of linguistic knowledge also underlie neural networks and connectionist models.

similarity relations exist in both form and meaning. Meaning is related to the function an object/sign has.<sup>16</sup> It is clear that signs may end up in the same paradigmatic set because their referents are seen to be similar in function. For instance, items like *apple*, *pear*, *banana* etc. do not form the set (sign) *fruit* so much on the basis of similarity of form/colour, but on the basis of similarity of function, i.e. they are all plucked, eaten, peeled, enjoyed in similar ways.<sup>17</sup> The analogical grid implies a close bond between the form and the function of a sign; it applies to all meaningful units, from the smallest morphemes, to complex words, but also to larger and more abstract (morpho)syntactic structures. Because form and meaning form a whole, a meaning change may affect the form, but change may also be driven by lexical items similar in form or by the more abstract formal requirements of the system. That form may drive meaning is nicely illustrated on a lexical level by Coates (1987), who shows how folk-etymological changes are often shaped by similarities in form.

Analogy is a basic force not only ontogenetically but also phylogenetically. Deacon (1997) shows that the grammatical, symbolic (i.e. abstract/arbitrary) system that became part of human language in the course of evolution was built up incrementally on the basis of the iconic and the indexical modes of thinking, guided by evolutionary old cognitive principles (i.e. the ability to see similarities and differences, the ability to categorize), which are also at work in other (non-linguistic) domains.

Iconic relationships are the most basic means by which things can be represented and it is the foundation on which all other forms of representation are built. What is important here is that iconicity depends on recognition, and recognition depends on the interpreter. When we interpret the world around us in terms of similarities and differences, we learn to see only differences which are functional or relevant, gradually ignoring non-functional ones. In other words, we don't learn and remember more than is absolutely necessary. This is what Hawkins (2004: 40) has called the principle of 'Minimize Forms':

*Minimizations in unique form-property pairings are accomplished by expanding the compatibility of certain forms with a wider range of properties [meanings]. Ambiguity, vagueness, and zero specification are efficient, inasmuch as they reduce the total number of forms that are needed in a language.*

---

<sup>16</sup> Note that the relation between the sign and its meaning/function is in itself indexical; children learn the meaning of linguistic signs because they are linked to a particular situation.

<sup>17</sup> Itkonen (2005) emphasizes that both function and form work analogically, and strengthen each other.

Hawkins continues that this minimization is connected with the frequency of the form and/or the processing ease of assigning a particular property to a reduced form. The ambiguity that arises is no problem since “[t]he multiple properties that are assignable to a given form can generally be reduced to a specific P[roperty] in actual language use by exploiting ‘context’ in various ways” (ibid: 41). For example, we learn to recognize phonemic and ignore phonetic distinctions in the course of language acquisition because the latter are not functional. In other words, it is more economic to ignore these differences.

What I am suggesting is that in the course of both language evolution and language learning, and hence also in language change, the same analogical reasoning keeps playing a role, whereby abstract items/structures gradually evolve from concrete (lexical) items constituting what Holyoak/Thagard (1995) have called ‘system mapping’.<sup>18</sup> System-mapping led to the evolution of grammar; it is still basically followed by children when they build up their grammar; and it guides language processing all through our lives. The exact path is not the same in all three domains because the input is different and keeps changing, but the same analogical principles are at work each time. An additional advantage of the analogical learning system is that there is only one system to begin with, i.e. a lexical one. It is therefore more parsimonious from an evolutionary point of view, and it better fits present neurological findings and the ideas developed about neural networks.<sup>19</sup>

In a frame like the above, analogy is both a mechanism and a cause. By means of analogy we may arrange linguistic signs (both concrete and abstract) into (other) paradigmatic sets, but it is also analogy that causes the learner to build up more abstract schemas, and to keep the number of these to a minimum (so it is a form of ‘grammar optimization’, but more local than suggested by Kiparsky (forthcoming), and always exemplar-based). In this learning model analogy is the primary force and not reanalysis. Reanalysis is what a linguist may see from the point of view of what changes in the system between generations or in the language output in the course of time, it is not

---

<sup>18</sup> Cf. Anttila (1977: 69) who writes that ‘all change mechanisms have an analogical ingredient.’

<sup>19</sup> This also enables us to see lexicalization and grammaticalization as basically the same process, but with different results based on different inputs (cf. Fischer 2007: 228-9). Goldberg (2001: 66) found that “knowledge of word meaning is not stored in the brain as a separate, compact module”. He also notes that the search for dissociations to establish the autonomy of a grammar module is fallacious in that for “every case of strong dissociations there are hundreds of cases of weak dissociations, where many functions are impaired together”(ibid: 56).

something that speakers actually do. Speakers do not reanalyse, they substitute one pattern holistically for another.<sup>20</sup>

Analogy is often seen as too loose, and therefore impractical or unworkable as a principle within a linguistic model. But indeed, it is not a principle of the system or a principle of language(-change), it is a faculty of language users. As Hofstadter (1995) emphasized, the conceptual mismatch that analogy represents, is in fact its strength: its flexibility keeps the system oiled. This is not to say that our analogizing capacities are not controlled. They are. The ‘looseness’ of analogy will be much constrained if one thinks of analogizing as taking place on different levels, and on concrete as well as abstract categories, all connected in tight networks. The possibilities are also constrained by the fact that the patterns and the paradigms are organized both semantically and structurally since each linguistic sign or token, be it single or complex, is, because of its binary nature, part of formal (sound-shape, structure, position) as well as semantic categories.<sup>21</sup>

This means that in order to discover how exactly analogy plays a role in grammaticalization processes or in change in general, one cannot concentrate on the development of one particular structure or (combination of) lexical item(s) only, one has to consider the change in terms of the network that the construction/item operates in. To get an idea how this works, it is useful to consider what happens in actual processing. Berg (1998) has looked at processing errors (and what causes them) as a way of determining the structure of the grammatical system.<sup>22</sup>

Berg makes a distinction between contextual and non-contextual errors. He shows how errors depend on “similarity constraints ‘elsewhere’” (1998: 173). Thus, an error like *cuff of coffee* is much more likely to occur than *hit the roop*. In both cases there is a [p]/[f] interchange, but in the first case the error is caused syntagmatically (by *coffee*),<sup>23</sup> and in the second

---

<sup>20</sup> If indeed we see constructions as “cognitive schemas of the same type as we find in other cognitive skills, that is, as *relatively automatized procedures for getting things done*” (Tomasello 2003b: 10, emphasis added), then it is also more likely that analogy rather than reanalysis is at work in processing, as the latter would entail some form of deliberation. Cf. also Bybee (2003: 155), who notes that speakers have no conscious access to grammatical knowledge, the latter “resembles procedural knowledge”.

<sup>21</sup> Cf. Berg (1998: 58ff.) for the description of such a network as a possible psycholinguistic model.

<sup>22</sup> Cf. Berg (1998: 165), who writes: “if the key to diachrony is to be found within synchrony – we should expect the patterns of language change to mirror the patterns in speakers’ and listeners’ spontaneous behaviour”.

<sup>23</sup> This is similar to priming, which interestingly enough is discussed by Jäger/Rosenbach (2008) as a possible cause for the unidirectionality of grammaticalization. However, as I will suggest below, it is

paradigmatically (i.e. [p] and [f] belong to the feature set of voiceless labials). Interestingly enough, with higher-level errors involving meaningful elements, *non-contextual* errors are much more likely to occur. Berg (1998: 165) gives the following example: *Muß sie es noch mal ticken – tippen?* ('Does she have to retype it?'), which he describes as an error that is "neutral with respect to the similarity scale, as there is nothing to compare [it] with" (ibid: 166), i.e. there is no [k] around in this case to cause the [p] in *tippen* to change to [k]. The interesting thing, however, is that both *ticken* and *tippen* are possible words in German. Moreover, semantically and formally they are very similar: both are verbs, they look alike phonetically, and both refer to a light, repetitive ticking sound.<sup>24</sup> Quite clearly, here, the error is of a paradigmatic kind, showing similarity on a deeper level of mental organization.

I would suggest that processing errors of the paradigmatic, non-contextual kind are more likely to be innovations that could result in actual change than contextual ones because the influence of paradigms in the grammar-system is likely to be stronger than the influence of context. The latter is bound to be variable, it being part of the actual discourse, while the former is much more stable, paradigms having become part of the system through learning and repeated use.<sup>25</sup> It has been shown in Analogical Modelling that changes in the morphological system are heavily constrained by the different paradigmatic sets that an item is part of (cf. Chapman/Skousen 2005).

Although such constraints are much more difficult to establish in the area of syntax (because the paradigmatic choices are so much wider; cf. footnote 10), promising work has been done here too showing that the development of constructions is not a linear affair (affecting only the particular construction under discussion) but 'starlike', influenced by other constructions that

---

probably the paradigmatic axis, rather than the syntagmatic one involved in priming, which plays the more important role in change.

<sup>24</sup> Onomatopoeia could play a role here, and even bilingualism or language contact if the speaker is familiar with another language like Dutch, where this verb would have a [k]. Recent research has shown that words of different languages are stored in a common lexicon and are accessed non-selectively, and that, although task-dependent, its first processing stages might remain unaffected by nonlinguistic contextual factors, cf. van Heuven et al. (2008) and studies referred to there.

<sup>25</sup> Of course when the context is fixed, as in grammaticalizing or lexicalizing constructions, then the construction itself can undergo change (phonetic attrition, bleaching, bonding etc.). But note that even here, a change involving a different category (e.g. from lexical verb to auxiliary) usually is due to paradigmatic similarity too, as argued above with respect to *going-to*. In a wider situational context, grammaticalization can also be driven by pragmatic inferencing in its first stage. For this to happen, the same situational context needs to be frequent. Note that here too analogy is involved because the inferencing depends on a comparison with previous situations in which the same structure occurred.

resemble them formally and/or semantically. DeSmet (2009, forthcoming) argues convincingly that certain cases that traditionally were seen as instances of reanalysis are better explained (in terms of the available data) as being driven by the presence of analogical forms elsewhere. Looking at the spread of the ‘new’ *for-NP-to-V* construction (with *for* functioning as complementizer and NP as subject) in English, he shows that this new construction became available because it was cast into the mould of an older but formally identical *for-NP-to-V* construction, where *for* was part of the infinitival marker *for... to* and the NP the *object* of the infinitive. The latter disappeared because the OV pattern itself was cast into the mould of the by then more regular Middle English clause-pattern, i.e. SVO, so that any NP before a verb came to be interpreted as subject rather than object. This explains better than the reanalysis story why the new construction doesn’t first appear as extraposed subject (as one would expect with reanalysis), and why there is an early predominance of passive infinitives. Another paradigmatic factor that facilitated the spread of the ‘new’ construction to more and more verbs was the analogy between the *for-NP* (in the subject-construction) that looked formally similar to the *for-NP* found as a prepositional object with the *same* verbs, causing the spread of the new subject-construction to other verbs taking a *for-PP*.

Other cases investigated show that grammaticalization doesn’t necessarily follow a gradual linear path but constitutes an abrupt process by analogy (Bisang 1998, Noel 2005). Fischer (2007: 274ff.) shows that in the cline from adverbial adjuncts to pragmatic markers in English, some of the pragmatic markers were attracted to the pattern directly via analogy, or via another pattern, that of reduced modal clauses. Similarly, she argues that in the development of English epistemic modals, there was no direct path from deontic to epistemic use. Epistemic meaning arose through functional and formal analogy with pairs of constructions like *he seems to be.../it seems that he...*, which enabled the *it must be that he...* to be replaced by *he must be....* These solutions are more commensurate with the philological facts and, as a further bonus, obliterate the problem that they do not neatly follow Lehmann’s parameters in terms of scope.

#### 4. Concluding remarks

Grammaticalization as a process only ‘exists’ on the language-output level. It may involve universal paths and may look unidirectional but this is not something intrinsic to the process on the speaker-listener level. As a process, it is an analyst’s generalization, a convenient summary but not something that

has actually ‘happened’ (cf. McMahon 2006: 173). Its apparent universality and directionality is caused by the fact that the lexical source items which are involved in it are (i) part of the basic vocabulary, (ii) as such are relatively frequent, (iii) are therefore likely to be phonetically and semantically reduced, which in turn (iv) makes them more eligible than other linguistic signs to function in abstract structural patterns. There is, however, no necessity about the development.

Language change can therefore not be *explained* in terms of grammaticalization. Grammaticalization occurs, and is often of a homogeneous ‘type’, especially when a form/construction through frequency has eroded so much that it becomes part of a drift,<sup>26</sup> but what ultimately decides whether a linguistic sign becomes part of a user’s grammatical system is whether it resembles in some ways (semantically, formally or both) an already existent category. Grammaticalization does not lead to new grammatical structures in any general sense (pace Meillet 1912, Bybee 2003, Kiparsky forthcoming, Traugott 2008: 154) except perhaps in cases of substratum or long-term contact, where new structures may enter through bilingualism or imperfect learning.<sup>27</sup> This may introduce genuinely new structures (but they would still be based on the analogy of contact/substrate structures), which may then be used as a pattern.<sup>28</sup> I have tried to show that reanalysis is an analyst’s concept; in terms of language *processing* it is based on our ability to analogize. This ability is steered by frequency, and it

---

<sup>26</sup> If we see this in terms of a complex adaptive system (cf. Ritt 2004), the unidirectionality of grammaticalization could be explained by the fact that in case of a choice, the most successful variant will be replicated, causing older variants to be lost (unless each variant can continue in separate paradigms as with *going to/gonna*). Since grammatical function items are more frequent, and their reduced phonology is neurally more economical (involving less cost), they will tend to replicate successfully, and hence become even more frequent, and more reduced over time.

<sup>27</sup> This would explain, for instance, why the future clitic *’ll* in English will not continue on the cline towards an affix, because the system of English does not allow inflectional *prefixes* on the main verb, and a tense-*suffix* on the preceding pronoun is unlikely to occur for semantic reasons (see Fischer 2007: 198, pace Hopper/Traugott 2003: 141).

<sup>28</sup> It would be interesting to learn more about how indeed new categories develop, such as for instance the determiner system in English. Is analogy (it being based on a similarity with an existing item/structure) still possible in such a case? McColl Millar (2000) suggests that contact may have played a role, Schlüter (2005) and Sommerer (2008) see influence from existing rhythmical patterns. In both cases analogy could still be at work.

includes analogical expansion, thus covering all the important factors mentioned under (iv) in section 1.

## References

- Anttila, Raimo. 1977. *Analogy*. The Hague: Mouton.
- Antilla, Raimo. 2003. "Analogy: the warp and woof of cognition". In Joseph, Brian D.; Janda, Richard D. (eds.). *The handbook of historical linguistics*. Oxford: Blackwell, 425-40.
- Berg, Thomas. 1998. *Linguistic structure and change. An explanation from language processing*. Oxford: Clarendon Press.
- Bisang, Walter. 1998. "Grammaticalization and language contact, constructions and positions". In Giacalone Ramat, A.; Hopper, Paul J. (eds.). *The limits of grammaticalization*. Amsterdam: Benjamins, 13-58.
- Brinton, Laurel J.; Traugott, Elizabeth C. 2005. *Lexicalization and language change*. Cambridge: Cambridge University Press.
- Bybee, Joan. 2003. "Cognitive processes in grammaticalization". In Tomasello, Michael (ed.). *The new psychology of language: cognitive and functional approaches to language structure*. Vol. 2. Mahwah: Erlbaum, 145-67.
- Chapman, Don; Skousen, Royal. 2005. "Analogical modeling and morphological change: the case of the adjectival negative prefix in English". *English Language and Linguistics* 9, 333-57.
- Chomsky, Noam; Halle, Morris. 1968. *The sound pattern of English*. New York: Harper and Row.
- Coates, Richard. 1987. "Pragmatic sources of analogical reformation". *Journal of Linguistics* 23, 319-40.
- Dąbrowska, Ewa. 2004. *Language, mind and brain: some psychological and neurological constraints on theories of grammar*. Washington D.C.: Georgetown University Press.
- Deacon, Terrence W. 1997. *The symbolic species. The co-evolution of language and the brain*. New York: Norton.
- Denison, David. 2006. "Category change and gradience in the determiner system". In van Kemenade, Ans; Los, Bettelou (eds.). *The handbook of the history of English*. London: Blackwell, 279-304.
- De Smet, Hendrik. 2009. "Analyzing reanalysis". *Lingua* 119, 1728-1755.
- De Smet, Hendrik. Forthcoming. "Grammatical interference: subject marker for and phrasal verb particle out". In Traugott, Elizabeth C.; Trousdale, G. (eds.). *Gradualness, gradience and grammaticalization*. Berlin: Mouton de Gruyter.
- Fischer, Olga. 1997. "The grammaticalisation of infinitival to in English compared with German and Dutch". In Hickey, Raymond.; Puppel, S. (eds.). *Language history and linguistic modelling. A Festschrift for Jacek Fisiak on his 60th birthday*. Berlin: Mouton de Gruyter, 265-80.
- Fischer, Olga. 2007. *Morphosyntactic change. Functional and formal perspectives*. Oxford: Oxford University Press.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.

- Goldberg, Elkhonon. 2001. *The executive brain. Frontal lobes and the civilized mind*. Oxford: Oxford University Press.
- Goldberg, Adele E. 2006. *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.
- Harris, Alice C.; Campbell, Lyle. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Haspelmath, Martin. 1989. "From purposive to infinitive - a universal path of grammaticization". *Folia Linguistica Historica* 10(1-2), 287-310.
- Haspelmath, Martin. 1998. "Does grammaticalization need reanalysis?". *Studies in Language* 22, 315-51.
- Haspelmath, Martin. 1999. "Why is grammaticalization irreversible?". *Linguistics* 37, 1043-68.
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Heine, Bernd; Claudi, Ulrike; Hünnemeyer, Friederike. 1991. "From cognition to grammar - evidence from African languages". In Traugott, Elizabeth C.; Heine, Bernd (eds.). *Approaches to grammaticalization I*. Amsterdam: Benjamins, 149-87.
- Heine, Bernd. 1994. "Grammaticalization as an explanatory parameter". In Pagliuca, William (ed.). *Perspectives on grammaticalization*. Amsterdam: Benjamins, 255-87.
- Himmelmann, Nikolaus P. 2004. "Lexicalization and grammaticalization: opposite or orthogonal?". In Bisang, W.; Himmelmann, N. P.; Wiemer, B. (eds.). *What makes grammaticalization? A look from its fringes and its components*. Berlin: Mouton de Gruyter, 21-42.
- Hofstadter, Douglas. 1995. *Fluid concepts and creative analogies*. New York: Basic Books.
- Holyoak, Keith J.; Thagard, Paul. 1995. *Mental leaps. Analogy in creative thought*. Cambridge Mass.: MIT Press.
- Hopper, Paul J. (1991). "On some principles of grammaticization". In Traugott, Elizabeth C.; Heine, Bernd (eds.). *Approaches to grammaticalization I*. Amsterdam: Benjamins, 17-35.
- Hopper, Paul. J.; Traugott, Elizabeth C. 2003. *Grammaticalization*. (2nd edition). Cambridge: Cambridge University Press.
- Itkonen, Esa. 2005. *Analogy as structure and process*. Amsterdam: Benjamins.
- Jäger, Gerhard; Rosenbach, Anette. 2008. "Priming and unidirectional language change". *Theoretical Linguistics* 34, 85-113.
- Janda, Richard D.; Joseph, Brian D. 2003. "On language, change, and language change - or, of history, linguistics, and historical linguistics". In Joseph, Brian D.; Janda, Richard D. (eds.). *The handbook of historical linguistics*. Oxford: Blackwell, 3-180.
- Joseph, Brian D. 2004. "Rescuing traditional historical linguistics from grammaticalization theory". In Fischer, Olga; Norde, Muriel; Perridon, Harry (eds.). *Up and down the cline - the nature of grammaticalization*. Amsterdam: Benjamins, 45-71.
- Kiparsky, Paul. Forthcoming. "Grammaticalization as optimization". In Jonas, D.; Whitman, J.; Garrett, A. (eds.). *Grammatical change: origins, nature, outcomes*. Oxford: Oxford University Press.  
(available on <http://www.stanford.edu/~kiparsky/Papers/yalegrammaticalization.pdf>).

- Lehmann, Christian. 1982 [1995]. *Thoughts on grammaticalization: a programmatic sketch*. Munich: Lincom Europa.
- Lehmann, Christian. 2004. "Theory and method in grammaticalization". In Diewald, Gabrielle (ed.). *Grammatikalisierung*, *Zeitschrift für germanistische Linguistik* 32, 152-87.
- McColl Millar, Robert. 2000. *System collapse, system rebirth: the demonstrative systems of English 900-1350 and the birth of the definite article*. Bern: Lang.
- McMahon, April. 1994. *Understanding language change*. Cambridge: Cambridge University Press.
- McMahon, April. 2006. "Restructuring Renaissance English". In Mugglestone, Linda (ed.). *The Oxford history of English*. Oxford: Oxford University Press, 147-77.
- Meillet, Antoine. 1912. "L'évolution des formes grammaticales". *Scientia* 12(26), 6.
- Mithun, Marianne. 1991. "The role of motivation in the emergence of grammatical categories: the grammaticization of subjects". In Traugott, Elizabeth C.; Heine, Bernd (eds.). *Approaches to grammaticalization I*. Amsterdam: Benjamins, 159-84.
- Noel, Dirk. 2005. "The productivity of a "source of information" construction: or, where grammaticalization theory and construction grammar meet". Paper given at the Fitigra Conference, Leuven, Belgium, February 2005.
- Peters, Ann M. 1985. "Language segmentation: operating principles for the perception and analysis of language" In Slobin, Dan I. (ed.). *The crosslinguistic study of language acquisition*. Vol. 2. Theoretical issues. Mahwah, N.J.: Erlbaum Associates, 1029-67.
- Plank, Frans. 1995. "Entgrammatisierung - Spiegelbild der Grammatisierung?". In Boretzky, N.; Dressler, W.; Orešnik, T.; Wurzel, W. U. (eds.). *Natürlichkeitstheorie und Sprachwandel*. Bochum: Brochmeyer, 199-219.
- Pulvermüller, Friedemann. 2002. *The neuroscience of language. On brain circuits of words and serial order*. Cambridge: Cambridge University Press.
- Ritt, Nikolaus. 2004. *Selfish sounds and linguistic evolution: a Darwinian approach to language change*. Cambridge: Cambridge University Press.
- Roberts, Ian; Roussou, Anna. 2003. *Syntactic change. A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.
- Rubba, Jo. 1994. "Grammaticization as semantic change: a case study of preposition development" In Pagliuca, William (ed.). *Perspectives on grammaticalization*. Amsterdam: Benjamins, 81-101.
- Schlüter, Julia. 2005. *Rhythmic grammar. The influence of rhythm on grammatical variation*. Berlin: Mouton de Gruyter.
- Slobin, Dan I. 1985. "Crosslinguistic evidence for the language-making capacity". In Slobin, Dan I. (ed.). *The crosslinguistic study of language acquisition*. Vol. 2. Theoretical issues. Mahwah, N.J.: Erlbaum Associates, 1158-1256.
- Sommerer, Lotte. 2008. "Noun phrase typology and the emergence of the definite article: analogy, accommodation and frequency effects". *Views* 17(1), 63-90 (available online at [http://www.univie.ac.at/Anglistik/views\\_0801b.pdf](http://www.univie.ac.at/Anglistik/views_0801b.pdf)).
- Tabor, Whitney; Traugott, Elizabeth C.. 1998. "Structural scope expansion and grammaticalization". In Giacalone Ramat, A.; Hopper, Paul J. (eds.). *The limits of grammaticalization*. Amsterdam: Benjamins, 229-72.
- Tomasello, Michael. 2003a. *Constructing a language. A usage-based theory of language acquisition*. Cambridge, Mass.: Harvard University Press.

- Tomasello, Michael. 2003b. "Introduction: some surprises for psychologists". In Tomasello, Michael (ed.). *The new psychology of language: cognitive and functional approaches to language structure*. Vol. 2. Mahwah: Erlbaum, 1-14.
- Traugott, Elizabeth C. 1982. "From propositional to textual and expressive meanings: some semantic-pragmatic aspects of grammaticalization". In Lehmann, W. P.; Malkiel, Y. (eds.). *Perspectives on historical linguistics*. Amsterdam: Benjamins, 245-71.
- Traugott, Elizabeth C. 1989. "On the rise of epistemic meanings in English: an example of subjectification in semantic change". *Language* 65, 31-55.
- Traugott, Elizabeth C.. 1995. "Subjectification in grammaticalisation". In Stein, Dieter; Wright, Suzan (eds.). *Subjectivity and subjectivisation in language*. Cambridge: Cambridge University Press, 31-54.
- Traugott, Elizabeth, C. 2008. "'All that he endeavoured to prove was ...': on the emergence of grammatical constructions in dialogal and dialogic contexts". In Cooper, R.; Kempson, Ruth (eds.). *Language in flux: dialogue coordination, language variation, change and evolution*. London: Kings College Publications, 143-77.
- Van Heuven, Walter J.B.; Schriefers, Herbert; Dijkstra, Ton; Hagoort, Peter. 2008. "Language conflict in the bilingual brain". *Cerebral Cortex* 18, 2706-16.
- Wanner, Dieter. 2006. *The power of analogy. An essay on historical linguistics*. Berlin: Mouton de Gruyter.

## *On the role of perception in the acquisition of the peach – pitch contrast by Polish learners of English*

*Kamil Kaźmierski, Vienna\**

### 1. Introduction

This paper tries to account for the difficulty that Polish learners of English have with producing the English /i:/ – /ɪ/ vowel contrast by investigating the hypothesis that the difficulty with producing the contrast has its roots in the inability to perceive it<sup>1</sup>. The above-mentioned difficulty is notoriously widespread among Polish learners. As any English teacher who has had experience with Poles will know, getting them to produce *leave* and *live*, *seat* and *sit*, *reach* and *rich*, *feel* and *fill*, or *eat* and *it* as pairs of differing words can be a struggle. What Polish learners have been observed to do is neutralize the contrast and use only one vowel in all the above mentioned words, a peripheral vowel similar in quality to English /i:/.

A substitution of this sort is hardly surprising as far as, for instance, French or Spanish learners of English are concerned. Their phonological systems have only one high front (unrounded) vowel, while English has two, traditionally transcribed as /i:/ and /ɪ/. Polish, unlike Spanish or French, but in common with English, however, has been thought of as having two contrasting (unrounded) sounds in the high front region of the vowel space. Aside from a peripheral /i/, it also has its lower and less advanced counterpart, traditionally transcribed as /i/. The bar across the symbol suggests that it is a central vowel. In acoustic terms, however, it is only slightly more central than English /ɪ/. It is acoustically a front (centralized) rather than a truly central vowel, as suggested by an acoustic test (Gonet 1993: 249), which found /i/ to

---

\* The author's e-mail for correspondence: [kamil.kazmierski@univie.ac.at](mailto:kamil.kazmierski@univie.ac.at).

<sup>1</sup> The present contribution is based on the author's MA thesis entitled *The 'pitch' – 'peach' problem - On the acquisition of the /i:/ – /ɪ/ contrast by Polish learners of English*, written at the English Department of Adam Mickiewicz University, Poznań, and supervised by prof. dr hab. Katarzyna Dziubalska-Kolaczyk

be situated between /i/ and /e/, rather than /i/ and /u/ in the acoustic space. With regard to the relationship between this Polish /i/ and English /i/, Gussmann (2007) states that “[/i/] in terms of its positioning on the vowel diagram, is very close to the English vowel of *bit*”, and that “the English *bit* and the Polish *byt* [bit] ‘existence’ sound remarkably similar” (Gussmann 2007: 1). On this reckoning, that is having two analogous phonological categories, Poles should be expected not to have any difficulty in perceiving the contrast. They should simply map the English contrast onto the similar vowel pair they already have in Polish. Somewhat intriguingly, they do not do so. What usually happens instead is that both English vowels are substituted in production by only one Polish vowel, namely /i/. Thus, the problem in production calls for an explanation. Following is the rationale behind considering perception of the two sounds responsible.

Despite the ease which Polish students of English could be predicted to have with distinguishing the two vowels, a typical students’ response to a teacher pointing out that the given two words are in fact pronounced with two different vowels is: “no, they sound the same”. The recurrence of such comments had contributed to my considering the inability to perceive the contrast responsible for the apparent failure to distinguish them in production. There are two reasons for which the hypothesis with perception might be worth pursuing. The first reason is that the two vowel contrasts are realized differently in the two languages. Aside from a quality difference, there is also a quantity difference between the vowels in English, while the feature of vowel length is absent from Polish. Additionally, the more peripheral vowel has a diphthongal character in English, a situation unlike in Polish, where both vowels are relatively pure. These differences might lead to Polish listeners relying on the ‘wrong’ acoustic cues (i.e. cues relevant for Polish, but not for English), and therefore failing to perceive the contrast. For example, in Polish, the vowel target is reached at the very beginning of its articulation, unlike in English, where a relatively long onglide stage is detectible. Thus, focusing on the initial portion of the vowel could result in misalignments. The second reason why Polish learners might fail to distinguish perceptually between the two vowels is the somewhat dubious phonological status of Polish /i/ and /i/. While most recent analyses treat them as separate phonemes (Jassem 2003, Gussmann 2007, Rubach 2007,), there are also some regarding them as allophones of a single phoneme (see Feldstein & Franks 2002), taking the view of Polish phonology in which the two sounds are in complementary distribution, with [i] surfacing after non-palatal consonants, and [i] elsewhere. The issue is not as clear-cut as it seems, since such an analysis relies on assigning phonemic status to palatalized

consonants, which in itself is not uncontroversial (for a comparison of the two approaches see Gussmann 2007).

Due to the reasons mentioned above, it seems legitimate to make the hypothesis that the contrast is not produced, because it is not perceived in the first place. To verify whether or not Poles differentiate perceptually between the English vowels in question, a perceptual test was conducted. The results of this test along with their potential implications are discussed in the present contribution.

## 2. The starting hypothesis

The possible explanation pursued in this paper is that both English vowels are perceptually assimilated by Polish learners to only one, rather than two Polish vowels, and that this perceptual neutralization of the contrast is then reflected in production.

The starting point for the study presented below is the widespread difficulty that Polish learners of English have been observed to have. Therefore, the subjects' production was not verified in the course of the experiment, although, admittedly, it would widen the perspective on the issue. In this paper, the notion that this difficulty with producing the contrast concerns the majority of learners is not called into question, and what is investigated is one possible explanation for this difficulty, i.e. that perceptual assimilation results in faulty production.

## 3. Approaches to the role of similarity in the acquisition of second language phonology

As phonemic contrasts of two languages are compared, the question of the role that similarity plays in second language acquisition has to be addressed. One take on this issue is represented by Robert Lado (1957), who, in the formulation of the Contrastive Analysis Hypothesis, observed that similarities between the target language and the native language of a learner lead to ease in acquiring an L2, and that differences lead to difficulties. This general idea, when applied to the phonological systems of Polish and English would make a prediction that the contrast under discussion here should be acquired easily. If we agree that in both languages there are two high front vowels, then two English vowels should be assimilated to two Polish vowels. This, however, is not the case. Lado already points out that a learner, when listening to L2 speech, hears L1 phonemes, not L2 phonemes (Lado 1957: 11). He does not,

however, go into detail about what happens when two foreign sounds are perceived as one native sound.

Research on categorical perception (Werker & Tees 1984) sheds light on the influence of the L1 on a person's perception of speech. Exposure to L1 speech leads to 'perceptual reorganization', i.e. to a situation where a listener pays attention only to those phonetic features of speech which are phonemically relevant in his or her language. Catherine Best's Perceptual Assimilation Model (PAM) is a model of speech perception that takes this into account and helps to conceptualize how speech sounds of an L2 are assimilated to L1 phonemes under the influence of language specific adjustments in speech perception. PAM also helps to conceptualize how this language specific way of perceiving speech can be un-learned in the process of second language acquisition.

According to PAM (Best 1994), when a learner is confronted with two L2 sounds, four things can happen:

- 1) Two foreign sounds can be assimilated to two native categories,
- 2) Two foreign sounds can be assimilated to a single native category, as equally good exemplars of it,
- 3) Two foreign sounds can be assimilated to a single native category, but one as a better exemplar and the other as worse,
- 4) The foreign sounds can be perceived as non-speech sounds (Best 1994: 191).

A comparison of the vowel systems of Polish and English allows potentially of any of the first three interpretations of how the English /i:/ – /ɪ/ contrast is perceived by Polish learners, the fourth one being excluded as there are no grounds to assume that either of the English vowels is not conceived of as a speech sound. A prediction that a situation like the one in point 1) seems feasible at first. That is to say, /i:/ should be assimilated to Polish /i/ and /ɪ/ should be assimilated to Polish /ɨ/. In such a scenario, the foreign contrast is realized by means of a native one, and no new phonemic categories have to be formed. However, the difficulties which the learners have been observed to have at the production level suggest that such a mapping does not take place. Instead, an instance of the assimilation of type 2) or 3) can be suspected.

If the starting point for a given learner for a specific contrast is like the one in point two or three, that is, when a learner's own phonology has only one phoneme where that of the target language has two, then the learner needs to establish a new category. The difference between the two types of assimilation is that a learner can make no distinction between two foreign sounds (type 2) or can evaluate one of the sounds as a good, while the other one as a deviant realization of the native category (type 3). If this is the

starting point for a learner, a new phonemic category has to be formed for one of the target sounds if the contrast is to be mastered. For a new category to be established, it is helpful for the new sound to be maximally different from the native one. So, on this level, it is actually dissimilarity between sounds of two languages that facilitates second language acquisition.

#### 4. The hypothesis of the present study

To relate the starting hypothesis behind the perceptual test to the categorization of the possible scenarios outlined above, I hypothesized that a situation like in point 3) above takes place, i.e. the two English phonemes are assimilated to a single Polish category, with English /i:/ being the better and English /ɪ/ the worse exemplar of Polish /i/. As a prerequisite for the hypothesis, an assumption has to be made that English /ɪ/ is more similar to Polish /i/ than to Polish /ɨ/. Although there are two vowels in Polish – /i, ɨ/ – and two vowels in English – /i:/, ɪ/ – in the same area of the acoustic space, my hypothesis is that both English vowels are sufficiently close to be assimilated, both acoustically and perceptually, to only one Polish vowel, namely to /i/. To verify whether there are acoustic grounds for such assimilation, acoustic data on the quality of the vowels was gathered through recordings, and the acoustics of the vowels were compared. To verify whether the perceptual assimilation actually takes place, a perceptual test was conducted.

If the two English vowels overlap acoustically with one Polish vowel, then there are grounds for them to be assimilated to a single category. If, in the next stage, the perceptual test confirms that such assimilation indeed takes place, then perception may be held responsible for problems with production. If not, however, then the reason for the faulty production of the contrast does not lie in the inability to perceive it.

#### 5. Acoustic properties of the four vowels

The vowels looked at here are the English vowels of *peach* /i:/ and *pit* /ɪ/ as well as the Polish vowels that learners could be expected to substitute for them, namely the vowel of *bić* ‘beat’ /i/ and the vowel of *być* ‘being’ /ɨ/.

All four vowels can be described as high and front, with /i:/ and /i/ being more peripheral than /ɪ/ and /ɨ/ (Jassem 2003: 105, Cruttenden 2008: 99). The English contrast is enhanced by the feature of length, absent from the Polish phonological system. The significance of this feature for the phonology of English is unclear. While certain authors see vowel length as the primary feature which maintains the contrast in English (cf. McMahon 2002: 72f), others acknowledge that only in comparable environments does length play a

crucial role in determining it (Roach 1991: 18). Yet others claim that length carries with it no contrastive weight whatsoever (Yavaş 2005: 79). Without making broad generalizations, it can be said that the difference of quantity is secondary to that of quality, as vowel length in English is strongly influenced by phonetic environments, to the extent that in certain environments the contrast is preserved by means of quality only, while the quantity is to a substantial degree dictated by the phonetic environment (Cruttenden 2008: 94f). A very clear example of the influence of phonetic context is the fact that the so-called ‘long’ vowel /i:/ can be shorter than the so-called ‘short’ vowel /ɪ/, if the former finds itself in a syllable closed by a voiceless consonant and the latter in a syllable closed by a voiced consonant (Kaźmierski 2009: 32). The quality distinction, on the other hand, remains to be expressed in these contexts. For example, the vowel of *beat* is shorter than the vowel of *bid*, so the feature of length cannot be used for determining the identity of the vowel. Qualitatively, however, the vowels are clearly distinct. Another supporting piece of evidence for the secondary importance of quantity comes from the results of an experiment involving a listening test (Clark & Hillenbrand 2003: 10ff), where portions of equal length (60ms) were extracted out of words containing either of the two vowels, /i:/ or /ɪ/. They were identified correctly at the rate of 79% and above. Also, a study involving a perceptual recognition test of vowels of digitally adjusted duration and quality (Bogacka 2003) suggests that native speakers of English rely on spectral cues (i.e. formant frequencies) for determining the contrast between high vowels /i:/ and /ɪ/, as well as /u:/ and /ʊ/ more so than on durational cues. Having this in mind, it seems safe to assume that though present in English, the feature of length is of secondary importance for maintaining the vowel contrast between /i:/ and /ɪ/, and so an acoustic comparison of the English and Polish vowels which takes into account quality only does have some currency.

In order to collect data on the acoustic properties of the vowels, recordings were made of native speakers of American English and of native speakers of Polish. Subjects read out lists of carrier sentences containing test words with either of the two vowels in question for the given language, i.e. /i:/ or /ɪ/ for the English subjects and /i/ or /ɨ/ for the Polish subjects (see detailed description below). The recordings were felt to be especially needed for Polish, as the data available was limited. Previous studies (Gonet 1993, Jassem 1999, Kleśta 1998), did not specifically focus on the two vowels, and so had only a few tokens containing them. Also, the number of subjects was small, which resulted in an overall small number of occurrences of the vowels relevant for the present study.

## 5.1 Recordings of American and Polish speakers – procedure

In the recordings of the English vowels, an overall number of 20 subjects participated, eleven of whom were female and nine male. They come from various regions of the United States, and belong to different age groups, with the age varying from 20 to 65. One subject was born in Germany, where she spent the first six years of her life, and four subjects had Polish ancestors in the third generation or earlier. In none of them, however, was a foreign (Polish or German) accent detectible.

The material used for the recordings took the shape of a list of token words embedded in carrier sentences. The tokens were one-syllable words of the CVC structure. The initial consonant was either a stop or a fricative, and the final consonant was a stop. The reason for the choice of words of homogenous structure, as well as for the avoidance of nasals, glides and liquids, was to enable the segmentation of the words for the purpose of vowel length measurements. Although length differences were not of primary interest for the English-Polish comparison, they were used for an intra-English comparison of the length relationship between vowels closed by voiced versus voiceless stops. The token words were embedded in a carrier sentence of the form *Say X one more time*. The reason for that was to ensure that the tokens receive the same degree of stress and be thus comparable. Apart from test words, a number of carrier sentences contained filler words in order to divert the subjects' attention away from the vowels tested (for a complete list of test words and filler words see Appendix A). The final list was randomized with the proviso that the first and last sentence was one containing a filler word (for the final list of sentences see Appendix B). The list took a following form:

*Say 'bat' one more time*  
*Say 'fit' one more time.*  
*Say 'bead' one more time.*

...

The subjects received instructions to read the list aloud at a moderate pace. Their performance of the task was recorded on a high quality portable recorder in the form of .wav files at the sampling frequency of 22 kHz. The analysis of the recordings was conducted with the help of the Praat speech analysis software (Boersma & Weenik 2007). Measurements of the first formant (F1) and the second formant (F2) were taken at a steady state or at midpoint where no steady state was discernible. Also, vowel duration was measured.

In the recordings of Polish vowels 19 native speakers of Polish participated, nine of whom were female and ten male. They were between the age of 19 and 26 at the time of the experiment. None of them was raised as a bilingual, although ten of the subjects are proficient speakers of English. Only one of them had spent a period longer than two months outside Poland, which was one year spent in Scotland.

The material used for the recordings took the shape of a list of token words embedded in carrier sentences. The tokens were two-syllable words of the CVCVC or CVCCV structure. Two-syllable rather than one-syllable words were used for Polish recordings since the number of Polish one-syllable minimal pairs for the contrast under investigation is insufficient for the goals of the present study. The consonants in the vicinity of the tested vowel were either stops or fricatives, and none of them was a velar, nasal, glide or liquid. The reason for such a choice of neighboring segments was to avoid their influence on the vowels studied. The token words were embedded in a carrier sentence of the form *Powiedz X jeszcze raz* 'Say X one more time'. In the end, the list took the following shape (for a complete list of test words and fillers see Appendix C, and for the final list of sentences see Appendix D):

*Powiedz „stok” jeszcze raz.*  
*Powiedz „widać” jeszcze raz.*  
*Powiedz „piski” jeszcze raz.*  
...

The procedure was analogous to the one employed for the recordings of American speakers. At the analysis stage, no vowel length measurements were taken in the case of Polish recordings, as this feature was not of interest here.

## 5.2 Comparison of acoustic properties of the Polish and the English vowels

An overview of the results of the recordings described above can be seen in Figure 1 (for females) and Figure 2 (for males), the results for the two genders being presented separately due to universal differences in the quality of vowels produced by women and men. The tokens are represented on a graph, where F1 values are represented on the vertical axis and F2 values on the horizontal axis. Polish and English vowels are allotted on the same graph to enable comparison.

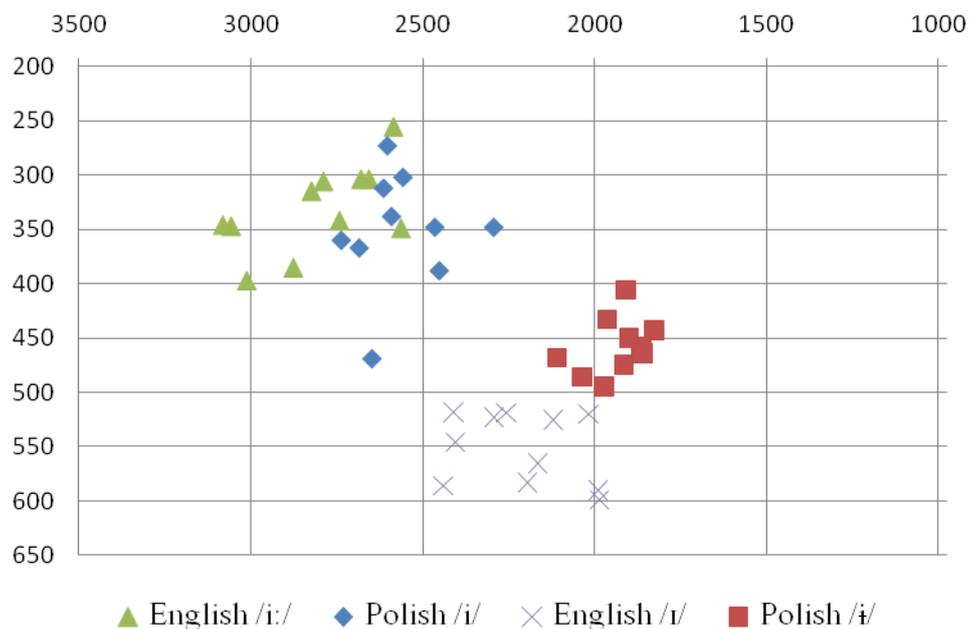


Figure 1: Polish and English high-front vowels – females

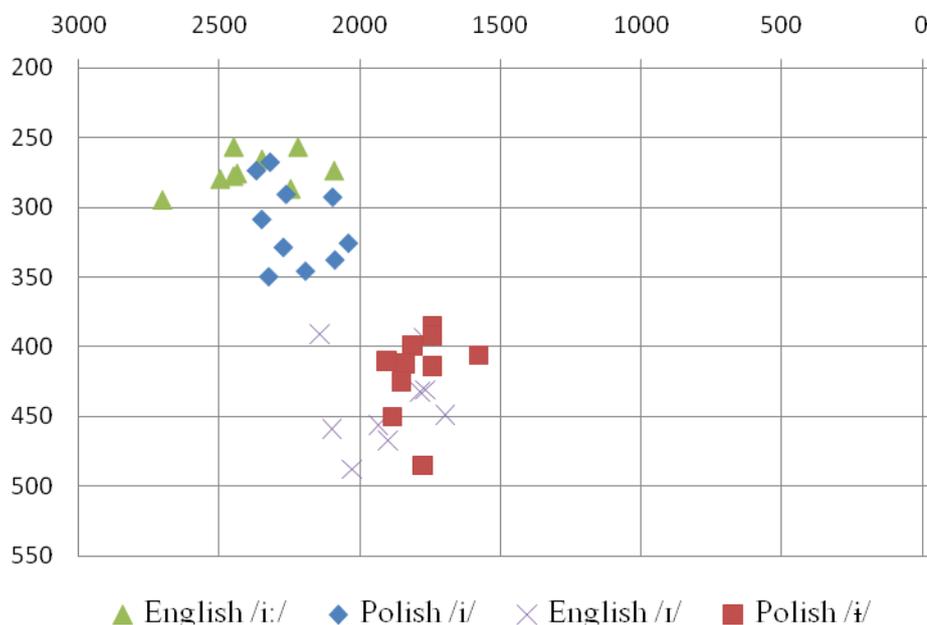
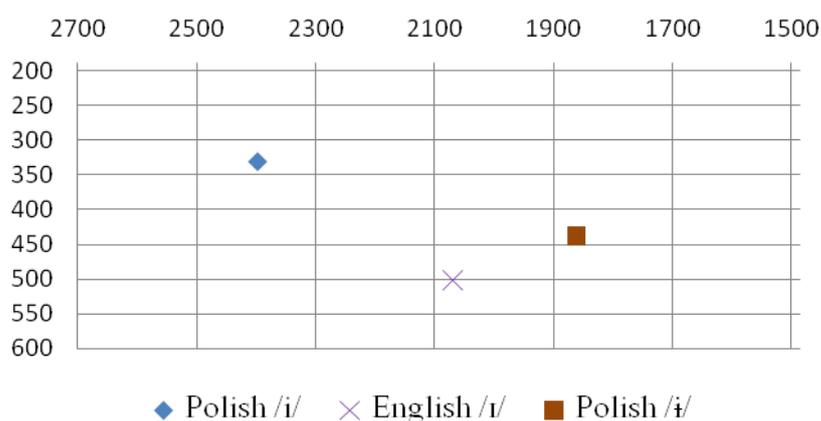


Figure 2: Polish and English high-front vowels – males

As can be seen, there is a good deal of acoustic overlap between /i:/ and /i/; and between /ɪ/ and /ɨ/ respectively. In the case of the first pair, the English vowel is more peripheral than the Polish one. In the case of the second pair,

the English vowel is situated slightly lower and more to the front in the acoustic space compared to the Polish one, especially for females. However, the distance between /ɪ/ and /i/ remains in all cases greater than the distance between /ɪ/ and /i/. In order to illustrate the relation of the English /ɪ/ to Polish /i/ and /i/, average distances between them are presented in Figure 3.



**Figure 3:** Average distance between /ɪ/, /i/ and /i/

The comparison presented above suggests that there is no reason, acoustically speaking, for the substitution of Polish /i/ for English /ɪ/, as there is a Polish vowel closer to the English one, namely /i/. Still, this comparison takes into account only the feature of vowel quality, represented by the first two formants of the vowel spectrum, and listeners could rely on additional cues for distinguishing between the vowels. Thus, a perceptual test is worthwhile to determine to which native categories are the English vowels assimilated when a learner is exposed to L2 speech.

## 6. Perception of the English vowel pair by Polish listeners

In the perceptual test, seven native speakers of Polish participated, six of whom were female and one male. Two of the female subjects took part only in one run of the test. All subjects were aged 16 at the time of the experiment, and they were all high school students learning two foreign languages, German and English. Polish speakers inexperienced in English would have represented a better sample, as their perception could not have been altered by exposure to English. However, due to technical difficulties subjects acquainted with some degree of English were used. Also, as their main second language was German, their English was hoped to exert little influence on their perception. In order to make sure that the test results be reliable, the test was run twice. The number of subjects was relatively small, but it is believed

that it was counterbalanced by a large number of tokens used in each test and by the fact that the test was run twice.

The above mentioned recordings of American English speakers were used as a source of samples for the perceptual test. Out of each recording, twelve tokens were extracted. Since there were twenty speakers altogether, the number of test words reached 240. These were then fed into a computer program that randomized them.

Subjects listened to successive words and were forced to label them as containing either of the two vowels, i.e. an 'i' or an 'y', where <i> and <y> are Polish spelling conventions for the qualitatively corresponding vowel pair in Polish, with <i> standing for the vowel /i/ roughly corresponding to English /i:/, and <y> standing for the vowel /ɨ/ roughly corresponding to English /ɪ/. Graphemes were used since subjects were not acquainted with the phonetic alphabet. Since it was the assimilation to native categories that was to be investigated, forcing the subjects to listen through the grid of their phonology was in keeping with the premises of the test. Additionally, the correspondence between spelling and pronunciation of the two vowels is nearly bi-unique in Polish, and so the graphemes could arguably be treated in this case as no less reliable representatives of the phonemes than the actual phonetic symbols would be. The grapheme 'i' nearly always refers to /i/ (except for cases where it does not stand for a vowel but acts as a marker of palatalization of the preceding consonant, e.g. *pies* 'dog' /pʲɛs/) and 'y' nearly always refers to /ɨ/ (except for the loanword *yeti* 'yeti' /'jɛti/, where the letter 'y' stands for the glide /j/).

In the course of the experiment, the majority of instances of the English /ɪ/ were assimilated to the Polish /ɨ/ and the majority of instances of the English /i:/ were assimilated to the Polish /i/. Thus, contrary to expectations, the listeners consistently discriminated between the two vowels. The greatest number of the anticipated 'misalignments', that is of assimilating the English /ɪ/ to Polish /i/, can be found in the case of speaker 4, who assimilated 25 per cent of the instances of /ɪ/ to the Polish /i/, with the percentage being smaller for the remaining subjects. The results of the first perceptual test can be seen in Figure 4 below.

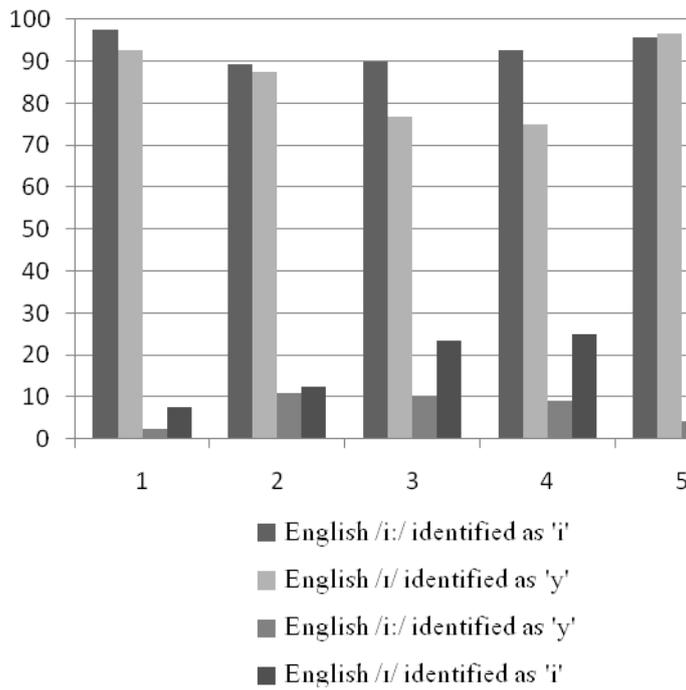


Figure 4: Assimilation of English vowels to Polish categories - test 1

At the second run of the test, the already small number of 'misalignments' fell further. In the case of speakers 1 and 5, it dropped to almost zero, and a drop is observable for all other speakers too. For the results of the second perceptual test see Figure 5.

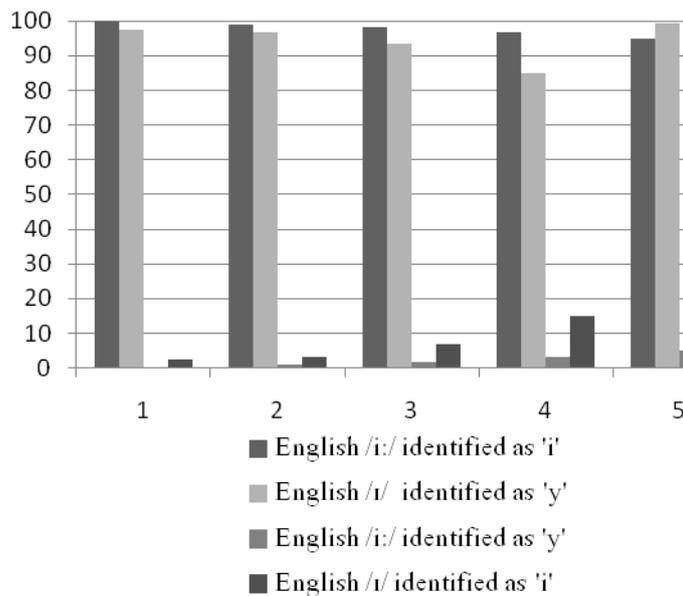


Figure 5: Assimilation of English vowels to Polish categories - test 2

Summing up, subjects could perceive the contrast in the test situation. This suggests that even though Poles generally have a problem with acquiring the English /i:/ - /ɪ/ contrast, they can perceive it. Admittedly, the very subjects who performed the perceptual test were not tested for production, and to advance any firm claims their production would also have to be verified. Yet, English was not their primary interest in foreign language study, and the production problem concerns even advanced Polish learners of English. Consequently, it could be assumed that their ability to perceive the contrast does not translate into the ability to produce it.

## 7. Conclusion

In this paper, I have tried to account for the situation where on comparing the phonological systems of two languages (Polish and English) we come to the conclusion that there is a systemic correspondence, and so learners of one of the languages should not have problems with acquiring this particular area of the second; however, actual encounters with Polish learners of English reveal that the contrast is not produced. An assumption made at the outset of this study was that although the two languages have two distinctive vowels in the same area of the vowel space, it is possible that their actual acoustic qualities are such that two phonemes of English are subsumed under only one category for a Polish listener, and that this perceptual assimilation of two foreign categories to a single native category can be held responsible for faulty production. However, measurements of the acoustic properties of the vowels in question showed that there are no grounds for such assimilation, and the results of the perceptual test confirm that it does not take place. Hence, the hypothesis has been falsified and the explanation for the faulty production has to be sought elsewhere.

A number of other paths of explanation could be pursued to try and account for the problem. The fact that the subjects were able to perceive the contrast, but cannot be assumed to produce it, poses two questions. The first question is general, namely how is it possible that a contrast can be perceived but not reproduced (and articulatory difficulty is disregarded here, as Poles do produce the contrast in their native language) and the second is specific, i.e. what is the reason for this particular problem that Polish learners of English face, if perception does not offer the answer. A tentative answer to the more specific question, namely that the distribution of the Polish phonemes hinders production in numerous environments, which then results in making false assumptions about English spelling at early stages in language learning is presented below.

One possible path is to investigate any potential effects of differences in distribution of the vowels in English and in Polish. It could be posited that when speaking English Poles are influenced by their L1 phonology. Although, arguably, it is a phoneme (Jassem 2003), the distribution of /i/ in Polish is restricted. For example, there is no contrast between the two vowels in certain environments, or the contrast is very rare (word initially, and after /k, g, l, dz/). When this restriction is applied to English words, it results in a number of mistakes, such as *it* \*/it/ (=eat), *live* \*/li:f/ (=leave, leaf), *list* \*/list/ (=least). Also, certain sequences, as /fi, ni, dzi/ are frequent in Polish, but only at morphological boundaries, and are virtually non-existent word-internally. Therefore, Polish learners may not construe them as valid sequences within English words, either; hence pronunciations such as *fill* \*/fil/ (=feel), *gin* \*/dzin/ (=Jean, gene), *animal* \*/'aɪnɪmə/ (with the coronal nasal /n/ palatalized to the palatal nasal /ɲ/ when followed by /i/, which is obligatory in Polish). It might be argued, then, that although when forced to decide, Poles do register the distinction (which means that speech perception is not blocking it), they transfer Polish distribution rules into English, and neutralize the contrast to /i/ in certain environments.

Another factor that might play a role is that of Polish spelling. As already mentioned, there is a near one-to-one correspondence in Polish, where the grapheme <i> nearly always stands for the sound /i/ and the grapheme <y> nearly always stands for the sound /ɨ/. The erroneous transfer of Polish spelling to pronunciation rules into English can be abetted by the fact that aside from 'looking like' /i/, grapheme <i> in English often appears in environments where only /i/ is allowed in Polish, for example *is, it, in, live*, or in environments, where /ɨ/ occurs extremely rarely, for example *give, kiss, fifty*. Thus, an explanation taking into account both distribution differences and spelling as reinforcing each other could be employed to account for production difficulties.

The role of phonotactics, however, would have to be verified through further research. Although there are certain consonant + /i/ combinations which are disallowed or dispreferred in Polish, there are still a number of others that are fully acceptable, and yet /ɨ/ is not produced in analogous consonant + /ɨ/ combinations in English anyway.

By means of an experiment involving perception and reproduction of the contrast in various phonological environments, conducted on subjects inexperienced in English, the validity of those tentative claims could be tested. Through such an experiment, it could be verified, whether in sequences used in Polish the contrast is reproduced more easily than in others.

Another suggestion for further research would be to investigate differences in the accuracy of the perception of the contrast between a forced-choice experiment and a more naturalistic setting. In the case of the experiment described above, the listeners knew that there was a contrast and had to seek relevant cues to establish it. When confronted with natural speech, however, the task of the listener is to decode what she hears, which does not necessarily presuppose that a sound heard at a given time has to be different from another one. Substantial differences in the rates of correct identification would point to the importance of discrimination exercises in listening training for acquiring the perception of L2 phonemes.

Although the hypothesis of the study has been falsified, the results of the experiments might have implications for teaching. Arguably, using the already present native contrast should be encouraged in the case of Polish learners. The avoidance of using it on the grounds that the quality of the Polish /i/ is not exactly the same as that of the English /ɪ/ seems counterproductive, as it has so far mostly lead learners to collapsing the contrast into one category. It seems that making the learners aware of the parallel between the English and Polish vowel contrasts and encouraging them to use the native Polish sound when they speak English would be helpful for them to keep the relevant word contrasts apart, even if the Polish sound is not a perfect equivalent of the English one. Still, exactly why this sound is thought of as not perfect and why there is resistance to using it in English remains a puzzling issue. After all, none of the remaining Polish vowels can be seen as a perfect counterpart of any of the English vowels, and still they are obviously substituted for the English vowel sounds.

## Appendix A: The list of test and filler words - English

Test words: Fit, Sit, Sick, Hid, Bid, Pit, Bit, Bead, Seat, Feat, Beat, Peat, Seek, Heed.

Filler words: fake, side, bat, sad, caught, code.

## Appendix B: The list of test sentences - English

Say 'bat' one more time

Say 'fit' one more time.

Say 'bead' one more time.

Say 'sit' one more time.

Say 'sick' one more time.

Say 'seat' one more time.

Say 'hid' one more time.

Say 'feat' one more time.

Say 'hid' one more time.  
Say 'bit' one more time.  
Say 'fake' one more time.  
Say 'heed' one more time.  
Say 'side' one more time.  
Say 'sick' one more time.  
Say 'bat' one more time.  
Say 'bid' one more time.  
Say 'beat' one more time.  
Say 'sad' one more time.  
Say 'side' one more time.  
Say 'feat' one more time.  
Say 'peat' one more time.  
Say 'pit' one more time.  
Say 'seek' one more time.  
Say 'beat' one more time.  
Say 'caught' one more time.  
Say 'sad' one more time.  
Say 'code' one more time.  
Say 'seek' one more time.  
Say 'caught' one more time.  
Say 'sit' one more time.  
Say 'bit' one more time.  
Say 'seat' one more time.  
Say 'bead' one more time.  
Say 'code' one more time.  
Say 'bid' one more time.  
Say 'heed' one more time.  
Say 'fit' one more time.  
Say 'peat' one more time.  
Say 'pit' one more time.  
Say 'fake' one more time.

## Appendix C: The list of test and filler words – Polish

Test words: widać, piski, pita, bity, biwak, Wisła, bywa, pyski, wyszła, byty, pytaj, wydam.

Filler words: jabłko, mięso, są, Irenka, stok.

## Appendix D: The list of test sentences - Polish

Powiedz „stok” jeszcze raz.  
Powiedz „widać” jeszcze raz.  
Powiedz „piski” jeszcze raz.  
Powiedz „bywa” jeszcze raz.  
Powiedz „pita” jeszcze raz.  
Powiedz „pyski” jeszcze raz.  
Powiedz „wyszła” jeszcze raz.

Powiedz „byty” jeszcze raz.  
 Powiedz „jabłko” jeszcze raz.  
 Powiedz „pytaj” jeszcze raz.  
 Powiedz „pita” jeszcze raz.  
 Powiedz „Irenka” jeszcze raz.  
 Powiedz „wydam” jeszcze raz.  
 Powiedz „stok” jeszcze raz.  
 Powiedz „bity” jeszcze raz.  
 Powiedz „piski” jeszcze raz.  
 Powiedz „Wisła” jeszcze raz.  
 Powiedz „wyszła” jeszcze raz.  
 Powiedz „jabłko” jeszcze raz.  
 Powiedz „Wisła” jeszcze raz.  
 Powiedz „mięso” jeszcze raz.  
 Powiedz „bywa” jeszcze raz.  
 Powiedz „byty” jeszcze raz.  
 Powiedz „biwak” jeszcze raz.  
 Powiedz „mięso” jeszcze raz.  
 Powiedz „wydam” jeszcze raz.  
 Powiedz „widać” jeszcze raz.  
 Powiedz „są” jeszcze raz.  
 Powiedz „biwak” jeszcze raz.  
 Powiedz „Irenka” jeszcze raz.  
 Powiedz „pyski” jeszcze raz.  
 Powiedz „bity” jeszcze raz.  
 Powiedz „pytaj” jeszcze raz.  
 Powiedz „są” jeszcze raz.

### **References:**

- Best, Catherine. 1994. “The emergence of native language phonological influences in infants: a perceptual assimilation model”. In Goodman, Judith C.; Nusbaum, Howard C. (eds.). *The Development of speech perception: The transition from speech sounds to spoken words*. London: MIT Press, 167-224.
- Boersma, Paul; Weenik, David. 2007. “Praat: doing phonetics by computer”. Version 5.1.04. <http://www.praat.org/> (5 Feb. 2009).
- Bogacka, Anna. 2003. “Perception of Second Language Speech Sounds”. MA thesis, Adam Mickiewicz University.
- Clark, Michael J.; James M. Hillenbrand. 2003. “Quality of American English Front Vowels Before /r/”. *Journal of the International Phonetic Association* 33(1), 1-16.
- Cruttenden, Alan. 2008. *Gimson’s Pronunciation of English*. London: Hodder Education.
- Feldstein, Ron F; Franks, Steven. 2002. *Polish*. München: LINCOM Europa.
- Gonet, Wiktor. 1993. “Próba Określenia Normy Wymowy Polskich Samogłosek Ustnych”. In Paprocka, Barbara (ed.). *Opuscula Logopaedica*. Lublin: Uniwersytet Marii Curie-Skłodowskiej, 232 – 253.
- Gussmann, Edmund. 2007. *The Phonology of Polish*. Oxford: Oxford University Press.

- Jassem, Wiktor. 1999. "Formants of the Polish Vowels as Phonemic and Speaker-Related Cues: Report on a Discrimination Analysis". *Speech and Language Technology* 3, 191 – 216.
- Jassem, Wiktor. 2003. "Polish". *Journal of the International Phonetic Association* 33, 103-107.
- Roach, Peter. 1991. *English Phonetics and Phonology*. Cambridge: Cambridge University Press.
- Rubach, Jerzy. 2007. "Feature Geometry from the Perspective of Polish, Russian and Ukrainian". *Linguistic Inquiry* 38(1), 85-138.
- Kaźmierski, Kamil. 2009. "The 'pitch' – 'peach' problem - On the acquisition of the /i:/ – /ɪ/ contrast by Polish learners of English". MA thesis, Adam Mickiewicz University.
- Kleśta, Janusz. 1998. "Linguistic and Paralinguistic Applications of Acoustic-Phonetic Database of Polish Vowels". *Speech and Language Technology* 2, 47 – 63.
- Lado, Robert. 1957. *Linguistics across cultures*. Ann Arbor: University of Michigan Press.
- McMahon, April M.S. 2002. *An Introduction to English Phonology*. Edinburgh: Edinburgh University Press.
- Werker, J. F.; Tees R. C. 1984. "Cross-language speech perception: evidence for perceptual reorganization during the first year of life". *Infant Behaviour and Development* 7, 49-63.
- Yavaş, Mehmet S. 2005. *Applied English Phonology*. Malden: Blackwell.

## ***“I know this word. (I think) It means ...” – Lexical competence and self-assessment in CLIL and traditional EFL classrooms***

*Eva Maria Seregély, Vienna\**

### 1. Introduction

*Content and Language Integrated Learning* (CLIL), the complete or partial use of a foreign language as the medium of instruction in subjects such as Biology, History or Maths, has gained increasing popularity in the European educational landscape in the last 15-20 years. In the Austrian context, the target language is English in the majority of cases, but might just as well be Croatian, Hungarian, Italian, French or any other language (compare Eurydice 2006: 5; 15-17).

The underlying concept of the approach is to provide students with a higher amount of foreign language input than would be possible in the traditional setting, that is, in the formal (E)FL classroom. What is crucial in this respect is that “CLIL is based on language acquisition rather than enforced learning” (Darn 2006), meaning that the development of linguistic skills should occur in a more natural way, resembling a child’s acquisition of their mother tongue (compare *CLIL Compendium* 2001).

Among the aspects most favourably influenced by CLIL is undoubtedly the learners’ lexicon. Through receiving content input in the foreign language, CLIL students acquire a greater and more profound knowledge of subject-specific terminologies than their traditional peers do (Dalton-Puffer 2007b: 5-6). The present paper aims to investigate whether the two groups also differ as regards general English vocabulary proficiency. Furthermore, the question will be considered as to how CLIL learners assess their own lexical competence and linguistic progress, compared to students attending conventional EFL classes.

---

\* The author’s e-mail for correspondence: [eva\\_seregely@gmx.at](mailto:eva_seregely@gmx.at).

First, the background to the present article is accounted for. After that, an attempt is made to define the concept of *lexical competence*, by introducing three different theoretical views. The succeeding section is then devoted to the research instrument, in terms of test design procedures and scoring system used. Next, the test results are presented, and complemented by some additional findings concerning background factors, before further implications about the impact and potential of CLIL teaching will be discussed.

## 2. Background

Both the theoretical considerations and the statistical analyses outlined in the present paper are based on research that I conducted in the course of my MA thesis in 2007 and 2008.<sup>1</sup> For this purpose, 33 students of grade eleven (i.e. third year of upper secondary school) at a Viennese grammar school were involved in an empirical study. Of these participants, 21 attended the so-called *Vienna Bilingual Schooling* (VBS) programme, a particular sub-form of CLIL instruction, intended not only for Austrian students who are especially gifted and particularly interested in English or other foreign languages, but also for students with English as their L1 or medium of communication, who already have some basic skills in the German language (“Bilinguale Schule”: 3 May 2009). The remaining 12 students were in a traditional class.<sup>2</sup>

As regards testing procedures and materials, my thesis is a replication of a large-scale, longitudinal study that was carried out in Sweden (Sylvén 2004). Instead of the originally three test rounds (conducted at different stages during the students’ upper secondary education), the test battery was reduced to a single sequence administered at 2-3 consecutive school days per class. In my attempt to examine as many different aspects of the students’ vocabulary skills as possible, I selected the following five lexical tasks (four types of tasks):

- a *self-report* test, in which the students had to estimate and demonstrate their knowledge of a series of lexical items,

---

<sup>1</sup> Cf. Seregély 2008; This thesis was written at the Department of English at the University of Vienna under the supervision of Ao. Univ.-Prof. Mag. Dr. Christiane Dalton-Puffer.

<sup>2</sup> For the distinction of the two test groups, the expressions CLIL and VBS vs. traditional are used interchangeably (cf. Seregély 2008: 30-36 for a comprehensive description of the context and the test groups.).

- a *words in context* test, in connection with a reading task, where the learners were asked to derive the meaning of a set of words from their use in a newspaper article,
- a *multiple choice* test, and
- two *cloze* tests, in which the participants had to complete the missing items in a selection of British and American English lexical phrases and idioms.<sup>3</sup>

Prior to the administration of these tests, sociolinguistic questionnaires were distributed to all students as well as the teachers of the CLIL group. These question sheets focused on background factors, such as the participants' native languages, time spent in English-speaking countries, leisure activities involving the use of English, as well as personal aims and motives, and general perceptions of the CLIL programme (see Seregély 2008: 39-42 and Appendix 2).

The main body of the present article will concentrate on the data obtained from the *self-report* test, as the outcomes of this specific test type have proven highly representative of the sequence as a whole.<sup>4</sup> The theoretical part, however, provides an outline of three different models of vocabulary knowledge, along with an overall evaluation of their relevance for the entire series of tests. As a starting point, some basic terminological issues will be clarified.

### 3. Defining the concept of *lexical competence*

In common, every day language, the term *word* is very often understood in the sense of “a single unit of language which means sth and can be spoken and written” (*Oxford advanced learner's dictionary* 2000: 1490). However, on closer examination of the relevant specialist literature, it becomes obvious that the definition is by no means as straightforward as that.<sup>5</sup> Thus, for the present discussion, it seems more adequate instead to speak of a *lexical item* or *lexeme*:

---

<sup>3</sup> For more details, cf. Seregély (2008: 36-38; 46 ff and Appendix 3).

<sup>4</sup> The correlation rate between the students' overall test scores and their results on the *self-report* test amounts to 0.95, symbolizing a particularly high degree of interrelation (in fact the highest compared to the remaining four tests; for more details, cf. Seregély 2008: 121).

<sup>5</sup> For instance, after comparing different expert and non-expert views, Singleton (1999: 8-38) arrives at the conclusion that now the *word* concept has become completely elusive. Similarly, Carter (1987: 3-32) argues that the notion of *word* is almost impossible to grasp.

**lexeme/lexical item** A separate unit of meaning, usually in the form of a word (e.g. ‘dog’), but also as a group of words (e.g. ‘dog in the manger’). (Widdowson 1996: 129)

This definition is also well in line with Cruse’s (1986, quoted in Sylvén 2004: 35) concept of the *lexical unit*:

*a lexical unit must be at least one semantic constituent*

*a lexical unit must be at least one word.*

In the following text, the terms *lexical item*, *lexical unit* and *lexeme* will be used interchangeably to refer to both, individual words, as well as phrasal expressions. The term *word* will only appear occasionally, when the focus is exclusively on an isolated, one-part linguistic component.

Very generally speaking then, a person’s or a learners’ *lexical competence* is their knowledge of and ability to use words and phrases (also: the *vocabulary*)<sup>6</sup> of a particular language. But what is actually involved in ‘knowing’ a lexical item? This issue has been debated extensively in the history of SLA research. Based on Sylvén (2004), I will focus on three principal approaches and their applicability to the present (and the original) work.

### 3.1 Richards (1976)

Richards’ paper on the role of vocabulary for language teaching and learning is not only one of the first, but also one of the most influential works in this specific field (Meara 1996: 1). Concerning the question as to what contributes to the development of lexical proficiency, he proposed the following eight basic assumptions:

1. *The native speaker [of a] language continues to expand his vocabulary in adulthood, whereas there is comparatively little development of syntax in adult life.*
2. *Knowing a word means knowing the degree of probability of encountering that word in speech or print. For many words, we also know the sort of words most likely to be found associated with the word.*
3. *Knowing a word implies knowing the limitations imposed on the use of the word according to variations of function and situation.*

---

<sup>6</sup> For a detailed discussion on the concept of *vocabulary*, see Seregély (2008: 23).

4. *Knowing a word means knowing the syntactic behaviour associated with that word.*
5. *Knowing a word entails knowledge of the underlying form of [a] word and the derivatives that can be made from it.*
6. *Knowing a word entails knowledge of the network of associations between that word and the other words in [a] language.*
7. *Knowing a word means knowing the semantic value of the word.*
8. *Knowing a word means knowing many of the different meanings associated with the word.*

(Richards 1976: 83, quoted from Meara 1996: 1-2)

According to Meara (1996: 2), Richards' model clearly reflects the theoretical concepts and research concerns that were predominant in the mid-1970s. The first assumption is based on studies about children's acquisition of their mother tongue and does not necessarily apply to the process of foreign language learning at later stages in life (compare Sylvén 2004: 36). Assumption 2 reflects the beginnings of research into Corpus Linguistics (Meara 1996: 2) and refers to the areas of word frequency and collocational patterns. The third assumption touches upon the fields of pragmatics and discourse analysis. The fourth derives from developments in syntactic theory. Assumption five concerns the category of morphology, and the sixth assumption covers the knowledge of synonyms, antonyms and other types of relationships between individual words (compare Meara 1996: 2; Sylvén 2004: 36). Finally, assumptions seven and eight include "the basic aspect of knowing a word, viz. what the word means" (Sylvén 2004: 36).

Richards does not cite any reasons for the internal hierarchy of these eight points. Yet, it seems strange that the two meaning-related categories should come last, while assumption one, which has little to do with the concept of lexical competence, and assumption two, whose relevance is also questionable in this respect, appear on the top of the list (see also Meara 1996: 3; Sylvén 2004: 36).

As regards the test battery used for my study (Seregély 2008), only the aspects of a word's "semantic value" (i.e. assumption eight) and its "different meanings" (assumption seven), as well as to some extent syntactic features (assumption 4) are taken into account (compare also Sylvén 2004: 36).

### 3.2 Carter (1987)

Eleven years after Richards, Carter (1987: 187) set out seven key requirements for knowing a word in a foreign language:

1. *It means knowing how to use it productively and having the ability to recall it for active use, although for some purposes only passive knowledge is necessary and some words for some users are only ever known passively.*
2. *It means knowing the likelihood of encountering the word in either spoken or written contexts or in both.*
3. *It means knowing the syntactic frames into which the word can be slotted and the underlying forms and derivations which can be made from it.*
4. *It means knowing the relations it contracts with other words in the language and with related words in an L1 as well.*
5. *It means perceiving the relative coreness of the word as well as its more marked pragmatic and discursal functions and its style levels.*
6. *It means knowing the different meanings associated with it and, often in a connected way, the range of its collocational patterns.*
7. *It means knowing words as part of or wholly as fixed expressions conveniently memorized to repeat – and adapt – as the occasion arises.*

Carter's list seems to be more relevant to the concept of lexical proficiency. Nonetheless, his and Richards' ideas partly overlap. Richards' assumption two unites Carter's aspects two and four. Conversely, Carter's third requirement covers both the fourth and the fifth assumption of Richards' model. Carter's aspect six combines the features expressed in Richards' assumptions six and eight, and Richards' third and Carter's fifth aspect also partly intersect (see Sylvén 2004: 37).

Carter's model actually reflects a particularly high level of lexical competence, as it defines the ability to understand and use lexical phrases and idioms (i.e. aspect seven) as one of the basic prerequisites for knowing a word (Sylvén 2004: 37). In my empirical study, this aspect has been considered within the framework of the two *cloze* tests. Moreover, this particular test type proves suitable for measuring word knowledge as specified under point one (see Seregély 2008: 36f; 58f for details). Carter's aspects three, four and five are covered in the *words in context* test (ibid: 54-55) as well as the *self-report* test (46 ff.), and, as will be shown in the succeeding section, in the latter, the first aspect is also taken into account. Aspect six is relevant to all four test types. On the other hand, the second aspect is not entailed in any of the testing procedures outlined above.

### 3.3 Henriksen (1999)

Henriksen's (1999) "Three dimensions of vocabulary development" are different from the models illustrated so far. Within this approach, lexical competence is not seen as a list of individual traits but as a set of different, yet interrelated, continua along which the learners' progress of knowledge can be described (Henriksen 1999: 304 and 315; Haastrup & Henriksen 2000: 222).

The first dimension, the *partial-precise knowledge dimension*, reflects "different levels of comprehension of the same lexical item" (Haastrup & Henriksen 2000: 222). It is an imaginary scale according to which the degree of internalization of such a lexeme into the learner's mental word store is measured (compare Haastrup & Henriksen 1998). On the *partial knowledge* end, we find the informant who recognizes the item and can make a rough guess as to its semantic value. On the contrary, the *precise* end of the scale is represented by the learner who does not only know the exact meaning of the word, is able to pronounce, translate and rephrase it correctly, but who can also "identify the domain<sup>7</sup> or provide word associations, and suggest other forms of the word" (Henriksen 1999: 305).

The second dimension, the so-called *depth of knowledge dimension*, comprises the areas of semantics, syntax and morphology (see Henriksen 1999: 305-306; Haastrup & Henriksen 2000: 222). In order to 'reach the top end of the scale', the informant must be familiar with the synonyms, antonyms and hyponyms (i.e. paradigmatic relations), as well as the collocational patterns (syntagmatic relations) of a lexeme. Furthermore, they must become acquainted with the syntactic and morphological features of the lexical item, and with any possible restrictions concerning its use (Henriksen 1999: 305-306).

As the name implies, dimension three, the *receptive-productive dimension*, reflects the scale ranging from mere word recognition to the productive use of an item in written or spoken discourse. Hence, this is the dimension "where the quality of the output is manifested" (Sylvén 2004: 39). Henriksen stresses that there is no clear-cut distinction between receptive and productive knowledge, but that the two merge and mix with each other, as the learner gradually gains greater familiarity (Henriksen 1999: 313), or - due to lack of practice - once again forgets how to use a particular word.

From the above description, it becomes obvious that the three dimensions are not to be seen as separate systems, but that they are actually closely

---

<sup>7</sup> Henriksen (1999) does not explain what she means by the concept of "domain". For more details on this particular issue, see Sylvén (2004: 38).

intertwined. Dimensions one and two both refer to different aspects of the semantization process (Henriksen 1999: 312). While they are directly linked with the acquisition of word knowledge,

*[d]imension 3 is essentially a control continuum that describes levels of access or use ability, which may be operationalized through different types of receptive and productive tasks (Henriksen 1999: 314).*

Sylvén (2004: 39) notes that in this connection, it would have been useful also to consider the pragmatic aspect, in a similar way as defined in Richards' assumption three (see above): "Being aware of the pragmatic properties of words is certainly important for the L2 learner" (Sylvén 2004: 39).

Apart from this minor shortcoming, however, Henriksen's model seems to provide the most useful framework for understanding the concept of lexical competence. In contrast to Richard's (1976) and Carter's (1987) lists of assumptions, it fully takes into account the different stages and problems involved in learning and 'knowing' a word. In the following section, it will be shown how the three dimensions are integrated into the design of the *self-report* test.

## 4. The self-report instrument<sup>8</sup>

### 4.1 Design and test administration

As indicated above, in the self-report test, the students were asked to estimate their level of knowledge of a set of target words. The format of this test type is based on the so-called Vocabulary Knowledge Scale (VKS), an instrument developed by Paribakht & Wesche (1997) in order to discern different stages in learners' acquisition of lexical competence. The basic idea behind the VKS is to join aspects of self-perception and linguistic performance on a 5-point scale to measure self-assessed and actual knowledge of specific lexemes in written form (Paribakht & Wesche 1997: 179). Table 1 illustrates the five answer categories and the level of knowledge they represent:

---

<sup>8</sup> See the appendix for an example of the *self-report* test.

	<b>Self-report categories</b>	<b>Level of knowledge</b>
A	I don't remember having seen this word before.	<i>total unfamiliarity</i> word and meaning unfamiliar
B	I have seen this word before, but I don't know what it means.	<i>partial unfamiliarity</i> word familiar, meaning unfamiliar
C	I have seen this word before and I <u>think</u> it means _____. (synonym or translation)	<i>partial recognition</i> word familiar, word meaning (synonym or translation) is guessed
D	I <u>know</u> this word. It means _____. (synonym or translation)	<i>total recognition</i> word familiar; correct meaning (synonym or translation) is given
E	I can use this word in a sentence: _____. (Write a sentence.) <i>If you do this section, please also do section D.</i>	<i>total recognition and correct word use</i> as in D In addition, the word is used with semantic and grammatical correctness in a sentence.

(Adapted from Paribakht & Wesche 1997: 179-180 and Sylvén 2004: 52)

Table 1: VKS scale - answer categories in the self-report test

Corresponding to the scale ratings, test scores range from 1-5 points per item. Answer categories A and B only require an indication of self-perceived word knowledge, which yields a score of 1 and 2 points respectively, while for any higher score to be awarded, a concrete proof of knowledge is necessary (Paribakht & Wesche 1997: 179-180). Incorrect responses in either of the categories C-E result in a score of 2. A score of 3 signifies that a correct synonym or translation has been written in one of the categories C or D. 4 points are allotted if the target word is used within a sentence which clearly displays the learner's understanding of its meaning in that particular context, but contains some grammatical error (e.g. the target word is used in another word class, misspelled or incorrectly conjugated as in 'catched' instead of 'caught'). For statements that demonstrate "both semantically and grammatically correct use of the target word, even if other parts of the sentence contain errors" (Paribakht & Wesche 1997: 180) the maximum score of 5 points is assigned.

This scheme allows for more detailed insights into the learners' actual level of lexical competence than a mere yes-no test could provide (Sylvén 2004: 53). In order to highlight any potential differences between the two groups regarding self-perceived and 'real' vocabulary knowledge, Sylvén (2004: 59) slightly altered the original scoring system: Synonyms or

translations written in category C, i.e. those that were **thought** to be known, only yielded a score of 3 if they were correct, and a score of 2 if they were incorrect. Four points were not awarded, except from category D (“I **know** this word ...”) onwards (For an illustration of this scheme, see Sylvén 2004: 59, Figure 4:3.). This modification was also adopted in the present context.

Just as in each of the other four tests in the battery, in the *self-report* test, 30 lexical items were under examination. A time limit of half an hour was set for this task (Seregély 2008: 43). Included in the test are words such as ‘a calamity’, ‘notwithstanding’ and ‘repulsive’. The selection of these items followed a highly complex scheme, based on the analysis of different word lists and corpora.<sup>9</sup> Apart from being categorized according to frequency bands, the *self-report* items are also fairly evenly distributed across word classes as well as the alphabet.

As regards Henriksen’s (1999) three-dimensional framework of knowledge, the *self-report* test clearly focuses on the *partial-precise* level, but proves equally suitable for measuring dimension 3 (cf. Sylvén 2004: 53-54). Answer category A signifies that the item is not familiar at all, while category B indicates that passive/receptive word knowledge exists. Columns C and D (see test sheet in the Appendix) represent different stages in the progress from the receptive to the productive end of the scale, and in column E, the learner has the opportunity to provide evidence of “his/her full productive control of the target word” (Sylvén 2004: 54). Dimension 2, the *depth of knowledge dimension*, is not covered within the current format of the test. However, Sylvén (2004: 54) points out that it might easily be included by adding a sixth answer category, in which the informants could demonstrate their knowledge of alternative forms and semantic features of each given lexeme.

## 4.2 Test scoring

For the most part, in the present test evaluation, the above-mentioned mode of scoring was employed. However, some of the participants’ responses were difficult to categorize according to the original principles. For these deviating answers, a new classification scheme had to be defined, on condition that the test scores remained within the original 1-5 point scale. Overall, twelve different types of problem sources could be discerned:

---

<sup>9</sup> Cf. Sylvén (2004: 62-68 and passim) for details.

- General problem sources:<sup>10</sup>

**Problem No. 1:** No answer is given at all.

**Description:** The student does not choose any of the five answer possibilities A-E but leaves an empty line next to the word under examination.

**Solution:** Since the original test design does not allow for any 0-point responses, and I could rule out the possibility that the participant does not know any English at all (and had therefore not understood the test instructions), this was interpreted as an instance of answer category A: “I do not remember having seen this word before.” For responses of this particular kind, a score of 1 point was awarded.

**Problem No. 2:** Demonstration of knowledge is missing.

**Description:** One of the columns C, D or E has been marked with a cross but no further information is given.

**Solution:** This was regarded as an instance of *partial unfamiliarity* (see Table 1 above). As in answer category B “I have seen this word before, but I do not know what it means.”, a score of 2 was assigned.

- Meaning-based problem sources:<sup>11</sup>

**Problem No. 3:** The correct word use is guessed.

**Description:** A sentence showing the correct use of the item in question is written in column E, but the corresponding synonym or translation appears in column C (instead of D).

**Solution:**

(a) If the meaning of the target word was guessed correctly, 5 points were awarded, as in:

(1)<sup>12</sup> **Item:** (to merge) (21)<sup>13</sup>

**Answer:** (C) ‘vermischen’

(E) ‘We get the result when we merge these two substances.’

[VBS.SR.19f]<sup>14</sup>

<sup>10</sup> This term refers to deviating answers resulting from careless mistakes or misinterpreting the test instructions.

<sup>11</sup> This category covers all those responses where the meaning of the target item is not understood or expressed in the required manner.

<sup>12</sup> Number of the linguistic example

<sup>13</sup> Number of the item in the *self-report* test

(b) If, on the other hand, an incorrect answer was given in column C, the student only received 2 points, as in:

- (2) **Item:** (to commence) (7)  
**Answer:** (C) 'carry on'  
 (E) 'Don't stop. Please commence.' [VBS.SR.13f]

**Problem No. 4:** The target word is used within a sentence but no further information is given.

**Description:** The student has written a sentence in column E, but no synonym or translation is given in column D.

**Solution:**

(a) If the sentence showed that the student was well aware of the word's meaning and its correct grammatical use, the maximum score (5 points) was awarded. An example of such an answer is:

- (3) **Item:** (to illuminate) (17)  
**Answer:** (E) 'The light illuminates the room.' [VBS.SR.11m]

(b) However, if the target word was used within a sentence which contained some inaccurate grammar (i.e. target word used within the wrong word class), as in

- (4) **Item:** (to disguise) (10)  
**Answer:** (E) 'You are the devil in disguise.' [VBS.SR.11m]

-- or which was ambiguous in terms of semantics, a score of 4 was assigned. An example of ambiguous word use is:

- (5) **Item:** (to commence) (7)  
**Answer:** (E) 'He commenced the race.' [VBS.SR.2f]

In the above context, both the correct meaning 'to begin', as well as the incorrect one 'to continue, to carry on', which was more frequently used by the students, might be possible.

This distinction between *correct* and *acceptable* answers was considered indispensable, since the mere fact that a learner knows how to use a particular

---

<sup>14</sup> In the following, this type of code will be used whenever an example is quoted from the students' test sheets.

lexical item within a sentence does not automatically imply that they are also familiar with its meaning. It might very well be the case that they have just “reproduce[d] the context in which they first met the word, or reproduce[d] a fixed expression which contains it” (Meara 1996: 6).

(c) On the other hand, sentences which clearly indicate that the student misinterpreted the meaning of the item under consideration were easily discerned. They yielded a score of 2, as in:

- (6) **Item:** (to illuminate) (17)  
**Answer:** (E) ‘We need to illuminate our dog.’ [VBS.SR.18m]

The 2-point solution for *Problem No. 4* is fairly similar to

**Problem No. 5:** Correct sentence but incorrect synonym/translation.

**Description:** Column E contains a sentence which shows the (seemingly) correct use of the target word; however, the corresponding translation or synonym in column D is incorrect.

**Solution:** Just as for all other incorrect answers, 2 points were assigned, as in:

- (7) **Item:** (in conclusion) (18)  
**Answer:** (D) ‘in Übereinstimmung’  
 (E) ‘In conclusion, you can say that ...’ [TRS.SR.3m]

**Problem No. 6:** Sentence in column E; both correct and incorrect synonym/translation written in column D.

**Solution:**

(a) If the sentence revealed a stronger ‘tendency’ towards the correct option, the student received 5 points, as in:

- (8) **Item:** (a refuge) (28)  
**Answer:** (D) ‘Flucht/Rückzugsort’  
 (E) ‘He hid in his refuge until they were gone.’ [TRS.SR.10f]

(b) On the contrary, sentences such as

- (9) **Item:** (an editorial) (11)  
**Answer:** (D) ‘Editorial/Inhaltsverzeichnis’  
 (E) ‘I still had to write the editorial for the book report.’ [VBS.SR.1f]

... where the content was more ‘inclined’ towards the wrong synonym or translation, yielded a score of 2.

Apart from these *meaning-based problem sources*, a series of answers were also problematic with regard to grammatical features:

- Grammatical problem sources:

**Problem No. 7:** Semantically correct, but grammatically inaccurate use of the target word within a sentence.

**Description:** Both columns D and E are filled in. Column D provides a correct synonym/translation of the target word. In column E, “the word is used within a sentence demonstrating the learner’s knowledge of its meaning in that context but with inaccurate grammar” (Paribakht & Wesche 1997: 180).

**Solution:** In accordance with the original scoring system, 4 points are awarded (compare Paribakht & Wesche 1997: 180 and Sylvén 2004: 59). An example of this specific answer type is:

(10) **Item:** (prediction) (26)  
**Answer:** (D) ‘Voraussagung  
 (E) ‘She predicted the future.’ [VBS.SR.7m]

The same also applied when the word-class confusion or grammatical error occurred in column D, whereas column E showed the semantically and grammatically correct use of the target word (= **Problem No. 8**), as in:

(11) **Item:** (a bargain) (4)  
**Answer:** (D) ‘to make profit’  
 (E) ‘The car I bought was a real bargain.’ [VBS.SR.13f]

However, if the word class was confused in both columns D and E, an additional point was subtracted, thus resulting in a score of 3 (= **Problem No. 9**). An example of such an answer is:

(12) **Item:** (to disguise) (10)  
**Answer:** (D) ‘a camouflage’  
 (E) ‘The military uses a disguise to not be seen.’ [VBS.SR.12m]

Consequently, grammatical errors in answer category C resulted in a score of 2, even if the answer was semantically acceptable (= **Problem No. 10**), as in:

- (13) **Item:** (hostile) (16)  
**Answer:** (C) 'Feind' [TRS.SR.2m]

The remaining two problem sources relate to aspects such as sentence complexity and collocations. They were subsumed into the category

- Stylistic problem sources:

**Problem No. 11:** The item is used with semantic and grammatical correctness in both columns D and E but the corresponding sentence is fragmentary.

**Solution:** Since the instruction for column E reads "... write a full sentence which shows how the word is typically used" (see Appendix), instead of the maximum score, only 4 points were assigned for a fragment such as:

- (14) **Item:** (a rate) (27)  
**Answer:** (D) 'Rate'  
 (E) 'unemployment rate' [TRS.SR.4f]

According to Sylvén's (e-mail: 13 March 2008) suggestion, stylistic errors and inappropriate collocations (= **Problem No. 12**), as in

- (15) **Item:** (a poll) (24)  
**Answer:** (D) 'Umfrage'  
 (E) 'The government started giving out a poll in which ...'  
 [TRS.SR.3m]

... were treated as those "other parts of the sentence" (Paribakht & Wesche 1997: 180) that were allowed to be erroneous without affecting the test score. For these instances, the maximum of 5 points was awarded.

Table 2 provides an overview of the *self-report* scoring applied in the present context and illustrates how the above-mentioned 'deviating' answers were integrated with the original categories:

Answer per column					Type	Answer category	Score
A	B	C	D	E			
X					regular	A	1
-	-	-	-	-	Problem No.1		
	X				regular	B	2
		X	-	-	Problem No.2		
			X	-			
		-	-	X			
		correct			regular, c	C correct	3
			wrong grammar	wrong grammar	Problem No.9		
		incorrect			regular, ic	C incorrect	2
		wrong grammar			Problem No.10		
			correct		regular, c	D correct	4
			correct	wrong grammar	Problem No.7		
			correct	fragment	Problem No.11		
			incorrect		regular, ic	D incorrect	2
			correct	correct	regular, c	E correct	5
		correct	-	correct	Problem No.3, c		
			-	correct	Problem No.4, c		
			inc./corr.	correct	Problem No.6, c		
			correct	wrong collocation	Problem No.12		
			-	wrong gr./ ambiguity	Problem No.4, a	E acceptable	4
			wrong grammar	correct	Problem No.8		
			incorrect	incorrect	regular, ic	E incorrect	2
		incorrect	-	correct	Problem No.3, ic		
			-	incorrect	Problem No.4, ic		
			incorrect	correct	Problem No.5		
			inc./corr.	incorrect	Problem No.6, ic		

c = correct, ic = incorrect, a = acceptable

Table 2: Summary - classification of self-report answers and corresponding scores

## 5. Discussion of results

### 5.1 Overall outcome

After all test sheets had been analysed and scored, the *self-report* test yielded the results illustrated in Figure 1:

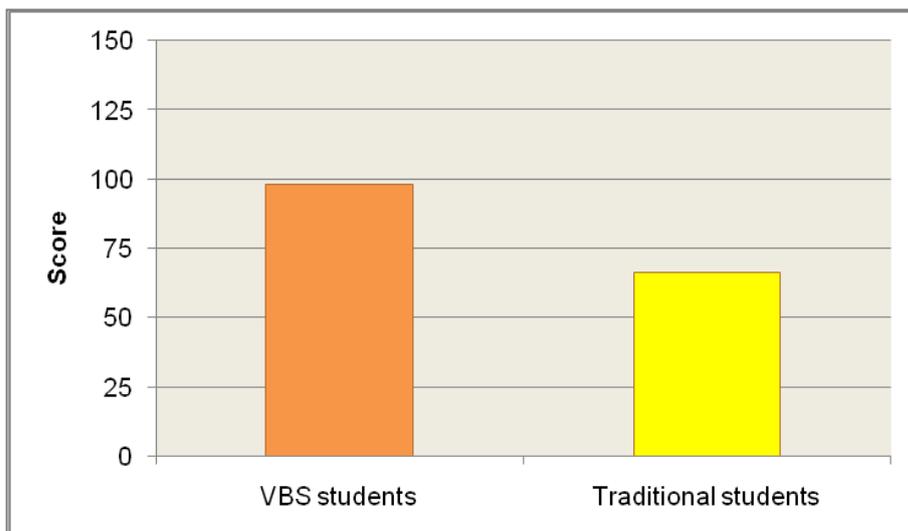


Figure 1: Mean score self-report test

The VBS (CLIL) group was clearly in the lead with a mean score of 97.90 of a maximum total of 150 points, compared to 66.42 points for the traditional group. The CLIL students' overwhelming superiority becomes even more noticeable if we consider the participants' individual scores.

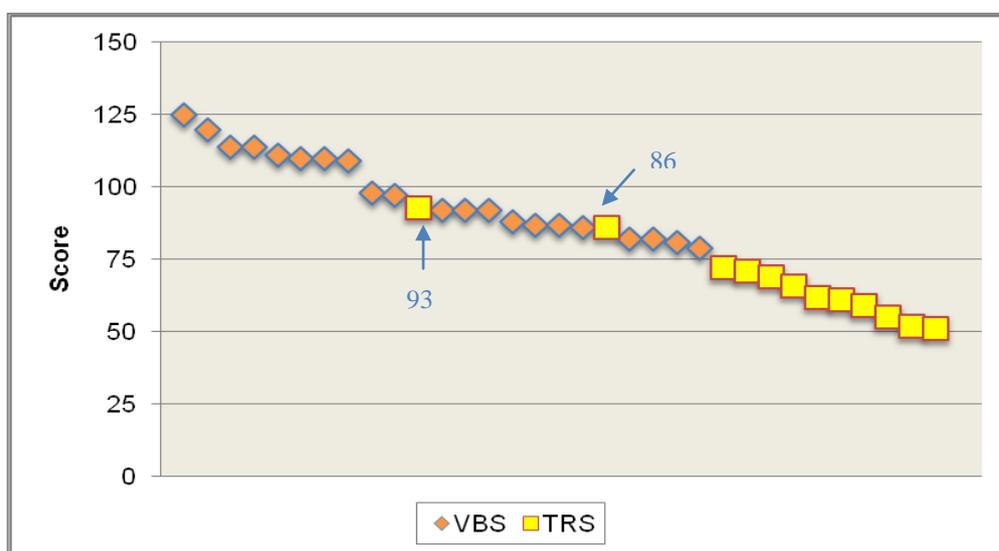


Figure 2: Self-report test – ranking of individual scores

As Figure 2 shows, the absolute results of the CLIL group extend from the scale's maximum of 125 to 79 points, while, apart from two exceptions, the traditional students' values lie between 72 and 51 points, which is the lowest score in the entire *self-report* range. In other words, only two of the EFL

learners managed to attain test results within the scope of their VBS peers'. The two scales intersect once, at the score level of 86, which is the second-best position on the traditional, as opposed to rank 17 (of 21), on the VBS scale.<sup>15</sup>

From the above figure, it becomes obvious that there is considerable variation not only between the mean scores of the two groups, but also between the values **within** each of the groups. It is therefore possible

*that the variation in mean scores is brought about by the variation between the individual subjects. In that case, the variation between the mean scores only reflects individual differences, not differences between the [teaching] methods<sup>16</sup> used. (Rietveld & van Hout 1993: 14)*

In order to preclude this worst-case scenario, first a paired F-test that is, an analysis of standard deviation/variance, has to be performed, which is based on the assumption of a normally distributed data set.

	test value $F_{pr}$	$\alpha$	F-value $F_{f_1, f_2; 1-\alpha/2}$
VBS vs. Traditional $n_1 = 21, n_2 = 12$	1,220884078	0,05	3,22614478
		0,01	4,85522025
		0,001	8,14150397

Table 3: Paired F-test: analysis of variance, results on the self-report test<sup>17</sup>

Table 3 illustrates that the test value  $F_{pr}$  clearly remains below the critical F-values for each of the significance levels 0.05, 0.01 and 0.001, meaning that the difference in variance between the two samples is not statistically significant but simply based on chance. Both groups can therefore be said to be fairly homogeneous.

Having established this essential condition, now the variation in mean scores can also be evaluated in statistical terms. For this purpose, an unpaired t-test is carried out:

<sup>15</sup> A similar finding also emerged in connection with the entire test battery (compare Seregély 2008: 61f.).

<sup>16</sup> i.e. CLIL vs. traditional EFL instruction

<sup>17</sup> For a definition of the above-mentioned variables and a detailed account on the statistical procedures employed in the present work, see Seregély (2008: 62-68).

VBS vs. Traditional $n_1 = 21, n_2 = 12$	mean difference $\bar{x}_1 - \bar{x}_2$	$s_d$	test value $t_{pr}$	$\alpha$	t-value $t_{n_1+n_2-2; 1-\alpha}$
				0,05	1,695518742
				0,01	2,45282418
	31,48809524	4,987541625	6,31334986	0,001	3,37489928

Table 4: Unpaired t-test, mean score self-report test

As the figures in Table 4 show, in contrast to the F-test, this time, the test value considerably exceeds the critical value, even at an alpha-level of as low as 0.001. Hence, there is a highly significant difference between the average test outcome of the CLIL group and the traditional group.<sup>18</sup> To sum up, the diverging results of the two groups are not based on any coincidental variations between individual participants, but reflect real differences, caused by the respective teaching approaches (and some other factors, closely related to those, as will be discussed at a later stage).

The following section reveals how this result is linked with the students' choice of response levels in the test.

## 5.2 Distribution of answers according to *self-report* categories

As illustrated in section 4, in the *self-report* test, the participants were offered five different answer levels, according to which they had to estimate and, if necessary, also demonstrate their knowledge of each lexical item under examination. The definitions and instructions for these levels are repeated below for convenience:

- A. "I do not remember having seen this word before."
- B. "I have seen this word before, but I do not know what it means."
- C. "I have seen this word and think it means (*write a synonym or translation*)"
- D. "I know this word. It means (*write a synonym or translation*)"
- E. "I can use this word in a sentence (*write a synonym or translation in column D and then go to column E and write a full sentence which shows how the word is typically used*)"

<sup>18</sup> Cf. Deutsche Gesellschaft für Qualität (1993: A 4, 11-12) for a comprehensive description of how to interpret test results at different significance levels.

Since quite a number of the students' answers were difficult to classify on the basis of this system, a complex new scheme was developed (see Table 2), which, while still adhering to the original 1-5 point scale (compare Sylvén 2004: 52-53; 59), also takes into account twelve different types of 'deviating' responses. Corresponding to these modifications, the following subcategories were defined (for full details, see Table 2):

- **A** (includes answers of problem type *No. 1*)
- **B** (+ answers of problem type *No. 2*)
- **C correct** (+ *Problem No. 9*)
- **C incorrect** (+ *Problem No. 10*)
- **D correct** (+ *Problems No. 7 and 11*)
- **D incorrect**
- **E correct** (including the correct 'versions' of *Problems No. 3, 4 and 6*, as well as *Problem No. 12*)
- **E acceptable** (meaning that the sentence in column E is correct but ambiguous, since no further information is given, or the answer in column D contains some inaccurate grammar, i.e. *Problems No. 4 and 8*)
- **E incorrect** (comprises the incorrect 'versions' of *Problems No. 3, 4, 5 and 6*)

Table 5 illustrates the distribution of test answers according to these categories (or levels) in the CLIL group and the traditional group:

Answer category	VBS		Traditional	
	N	%	N	%
A	58	9%	91	25%
B	140	22%	143	40%
C correct	15	2%	18	5%
C incorrect	27	4%	34	9%
D correct	33	5%	28	8%
D incorrect	16	3%	11	3%
E correct	243	39%	31	9%
E acceptable	19	3%	0	0%
E incorrect	79	13%	4	1%
TOTAL	630	100%	360	100%

Table 5: Distribution of answers according to self-report categories

If we simply consider the figures listed in the first line, the above-mentioned differences in mean scores are easy to understand. Answer level A, symbolizing total unfamiliarity with the item(s) under consideration (compare Table 1), accounts for only 9 percent of all responses among the VBS students, whereas in the traditional group, as many as 25 percent of the answers were categorized in the same way. On the contrary, option E,

representing the highest level of lexical knowledge, was more frequently chosen in the bilingual class. Here, approximately 39 percent of all responses (compared to 9 percent in the traditional group) were classified as *E correct*, 13 percent (vs. 1 percent in the traditional group) were considered to be (*E incorrect*), and 3 percent (0 percent) fell into the new category *E acceptable*. Level B, on the other hand, was especially popular in the traditional class, where 40 percent of all *self-report* responses (as opposed to 22 percent in the VBS class) were recorded in this second answer column.

As regards level C, “I have seen this word and **think** it means ...”, once again, the regular students were ‘in the lead’, with 5 percent of all answers having been categorized as *C correct* (2 percent in the VBS group) and 9 percent (4 percent) bearing the label *C incorrect*. These figures might indicate that, in accordance with Sylvén’s (2004: 185) initial assumption, the traditional learners were indeed less self-confident about their knowledge of particular test items than their VBS peers were.

However, if we compare the related proportions for the *actual* knowledge-level D (“I **know** this word. It means ...”), the differences between the two groups turn out to be less striking, but rather contradictory. While in the traditional class, in total, 11 percent of the students’ responses were organized into one of the D-categories (8 percent *D correct*, 3 percent *D incorrect*), the corresponding value for the VBS class amounts to only 8 percent (5 percent *D correct*, 3 percent *D incorrect*).

At first glance, this seems to mean that “[t]he CLIL student is not necessarily more self-confident with regard to English vocabulary than the control student” (Sylvén 2004: 186). Yet, this conclusion appears completely absurd, as it has already been shown that, instead of providing only a simple translation or synonym, “[t]he CLIL students who believe they know a word [most typically] go all the way to the E-level” (Sylvén 2004: 186) in order to demonstrate their knowledge. Needless to say, this tendency is more representative of their (high) self-confidence than the outcomes in any other of the five *self-report* categories.

Overall then, if the emphasis is placed exclusively on those answers levels where a concrete proof of knowledge is required, Table 5 reveals that synonyms, translations and sentence fragments (i.e. all kinds of responses subsumed into either of the categories C or D) were preferred by the traditional students (25 percent vs. 14 percent in the VBS group). On the contrary, in the VBS sample, 55 percent of the *self-report* responses (10 percent in the traditional sample) were given in full sentence form, with or without any further specifications in column D. In fact, one of the CLIL

learners' most popular E-level answer types was *Problem No. 4* (see section 4.2.).

The insight that the VBS students were more likely to use a particular test item within its appropriate context (even if they were not certain about its exact meaning, see also Meara's (1996: 6) quote above), while the traditional participants typically demonstrated their knowledge in forms that resemble an entry in a dictionary or vocabulary list is also quite interesting from a wider perspective. To me it seems that to a certain extent, this answering pattern actually reflects the two teaching approaches per se, more precisely, the way in which (general) target language items are presented and learned/acquired in the CLIL compared to the formal EFL classroom.<sup>19</sup> As will be briefly touched upon in the following section, this view has found additional support in the analysis of selected background factors evaluated in Seregély (2008) by means of student questionnaires.

## 6. The wider perspective

Closely connected to the participants' perception of their knowledge of the items in the *self-report* test was also the question as to how they estimated their own linguistic progress in English in the course of secondary education. For this purpose, in the sociolinguistic questionnaire, the learners were asked to assess their development in the fields of lexis, grammar, speaking, writing and understanding written and spoken discourse according to the Austrian marking system (compare Seregély 2008: 39 ff.). Without reproducing the exact details of the corresponding discussion (cf. Seregély 2008: 138 ff.), one result can be stated summarily: The VBS students rated themselves considerably higher in all six areas than their traditional colleagues did. This finding exactly corresponds to what one might expect in view of the "motivating, low-anxiety communicative atmosphere" (Dalton-Puffer 2007a: 174) that is said to be a characteristic feature of CLIL classrooms.

In both test groups, the greatest improvement was recorded in the category *listening comprehension* or "Understanding spoken English" (Sylvén 2004: 184), followed by *reading comprehension*. *Writing* in English ranked third, and *vocabulary* development occupied the fifth place. Finally, *grammar* turned out to be the area with regard to which all of the participants felt that they had achieved the least progress.<sup>20</sup>

---

<sup>19</sup> For more details on the role of vocabulary in different instructional settings, cf., e.g., Seregély (2008: 8-21) and Sylvén (2004: 25-35).

<sup>20</sup> Cf. Seregély (2008: 139-140) for the exact numerical values.

As far as the CLIL group is concerned, this ranking, once again, hints at the ‘priorities’ of the method itself, that is, in particular, “more language exposure and input” (Sylvén 2004: 4), as well as the basic rule that

*Fluency is more important than accuracy and errors are a natural part of language learning. Learners develop fluency in English by using English to communicate for a variety of purposes [and r]eading is the essential skill. (Darn 2006)*

For an interpretation of the traditional group’s result, however, more detailed insights into the actual EFL classroom procedures would be required.

## 7. Conclusion

The results of the study outlined in the present paper have shown that students receiving CLIL instruction are superior in lexical competence to learners taught in the traditional way. Moreover, the representatives of the two groups differ in terms of how they estimate and demonstrate their knowledge of general English vocabulary. While the VBS subjects do not hesitate to use the test items within more or less complex sentences, their control group peers most typically provide one- or two-word answers, very often combined with an expression of uncertainty concerning the meaning of the respective target lexeme (“I **think** it means ...”). Similar differences in self-assessment have also been detected in connection with the participants’ evaluation of their general linguistic progress over a longer period of time.

It may therefore be concluded that, amongst other benefits, CLIL favourably affects the learners’ confidence, motivation and perception of their own efficacy (compare, e.g., Marsh & Langé 2000; *CLIL Compendium* 2001; *Mål i mun* 2002 referred to in Sylvén 2004; Mewald 2004; Dalton-Puffer 2007a; Darn 2006; Abendroth-Timmer 2007; Lasagabaster & Sierra 2009). Not surprisingly, this in turn also enhances

*the language acquisition process at large, whereas students whose attitudes are less positive and who lack motivation will have greater difficulties improving their lexical proficiency (Sylvén 2004: 226).*

It is difficult to determine whether it is the CLIL learners’ positive self-assessment that leads to their high-level lexical performance, or their high-level lexical performance boosting their self-confidence, or whether both are actually dependent on other factors. However, what is unquestionable is that in the present context, *Content and Language Integrated Learning* has fostered both the psychological as well as the linguistic aspect. Thus, to round off the present discussion, I simply want to cite a VBS teacher’s statement on the prospective future of CLIL:

*It will spread. – The advantages are obvious I think.*

[QR.NST.1m]

## Appendix

### Self-report test

The following test sheets were originally designed for a study conducted in Sweden (Sylvén 2004) and have kindly been provided by the researcher for the present replication.

NAME: \_\_\_\_\_  
 CLASS: \_\_\_\_\_

**1. SELF REPORT**

Use the column which best corresponds to your knowledge of the word. The columns are the following:

- A. I do not remember having seen this word before.
- B. I have seen this word before, but I do not know what it means.
- C. I have seen this word and think it means (write a synonym or translation)
- D. I know this word. It means (write a synonym or translation)
- E. I can use this word in a sentence (write a synonym or translation in column D and then go to column E and write a full sentence which shows how the word is typically used)

	A	B	C	D	E
1. Adjacent					
2. To adjust					
3. An attire					
4. A bargain					
5. A calamity					
6. A clue					
7. To commence					
8. Crucial					
9. To devalue					
10. To disguise					

Do not use this column

--	--	--	--	--

--	--	--	--	--

Do not use  
this column!

	A	B	C	D	E
11. An editorial					
12. To emerge					
13. To exaggerate					
14. A fraud					
15. Furthermore					
16. Hostile					
17. To illuminate					
18. In conclusion					
19. Incidentally					
20. Ingenuity					
21. To merge					
22. A mortgage					
23. Notwithstanding					
24. A poll					
25. To precede					
26. A prediction					
27. A rate					
28. A refuge					
29. To rely					
30. Repulsive					

--	--	--	--

--	--	--	--

--	--	--	--

--	--	--	--

## References

- Abendroth-Timmer, Dagmar. 2007. *Akzeptanz und Motivation: empirische Ansätze zur Erforschung des unterrichtlichen Einsatzes von bilingualen und mehrsprachigen Modulen*. Frankfurt: Peter Lang.
- Carter, Ronald. 1987. *Vocabulary – applied linguistic perspectives*. London: Allen & Unwin.
- CLIL-Compendium. 2001. <<http://www.clilcompendium.com/clilcontacts.htm>>.
- Cruse, Alan. 1986. *Lexical semantics*. Cambridge: Cambridge University Press.
- Dalton-Puffer, Christiane. 2007a. *Immersion classrooms as language learning environment: applied linguistics perspectives*. Habilitationsschrift. University of Vienna.
- Dalton-Puffer, Christiane. 2007b. "Outcomes and processes in Content and Language Integrated Learning (CLIL): current research from Europe". In Delanoy, Werner; Volkman, Laurenz (eds.). *Further perspectives for English Language Teaching*. Heidelberg: Carl Winter, 1-15.
- Darn, Steve. 2006. "Content and Language Integrated Learning". *British Council BBC: Teaching English*. <<http://www.teachingenglish.org.uk/think/articles/content-language-integrated-learning>> (4 May 2009).
- Deutsche Gesellschaft für Qualität e.V. 1993. *Gesamtlehrgang Teil II/A*. (2. Ausgabe). Frankfurt: Deutsche Gesellschaft für Qualität e.V.
- Eurydice. 2006. *Content and language Integrated learning (CLIL) at school in Europe: Austria, national description – 2004/05*. <[http://eacea.ec.europa.eu/ressources/eurydice/pdf/070DN/070\\_AT\\_EN.pdf](http://eacea.ec.europa.eu/ressources/eurydice/pdf/070DN/070_AT_EN.pdf)> (4 May 2009).
- GRg 23/VBS. 2008. "Bilinguale Schule – Bilingual Schooling". *Gymnasium Draschestraße Vienna Bilingual Schooling*. <<http://www.grg23vbs.ac.at/home.html>> (3 May 2009).
- Haastrup, Kirsten; Henriksen, Birgit. 1998. "Vocabulary acquisition: from partial to precise understanding". In Haastrup, Kirsten; Viberg, Åke (eds.). *Perspectives in lexical acquisition in a second language*. Lund: Lund University Press, 97-123.
- Haastrup, Kirsten; Henriksen, Birgit. 2000. "Vocabulary acquisition: acquiring depth of knowledge through network building". *International Journal of Applied Linguistics* 10:2, 221-240.
- Henriksen, Birgit. 1999. "Three dimensions of vocabulary development". *Studies in Second Language Acquisition* 21:2, 303-317.
- Lasagabaster, David; Sierra, Juan Manuel. 2009. "Language attitudes in CLIL and traditional EFL classes". *International CLIL Research Journal* 1:2, 4-17. [online: <<http://www.icrj.eu/index.php?vol=12&page=73>> (10 May 2009)]
- Mål i mun. 2002. SOU 2002:27. Stockholm: Fritzes.
- Marsh, David; Langé, Gisella (eds.). 2000. *Using languages to learn and learning to use languages*. Jyväskylä: University of Jyväskylä.
- Meara, Paul. 1996. "The vocabulary knowledge framework". *The Lognostics Virtual Library*. <<http://www.lognostics.co.uk/vlibrary/meara1996c.pdf>> (3 May 2009).
- Mewald, Claudia. 2004. *Paradise lost and found: a case study of content based foreign language education in Lower Austria*. Unpublished PhD thesis. Norwich: Department of Education and Professional Development, University of East Anglia.

*Oxford advanced learner's dictionary of current English*. 2000. (6<sup>th</sup> edition). Oxford: Oxford University Press.

Paribakht, T. Sima; Wesche, Marjorie. 1997. "Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition". In Coady, James; Huckin, Thomas (eds.). *Second language vocabulary acquisition*. Cambridge: Cambridge University Press, 174-200.

Richards, Jack C. 1976. "The role of vocabulary teaching". *TESOL Quarterly* 10:1, 77-89.

Rietveld, Toni; van Hout, Roeland. 1993. *Statistical techniques for the study of language and language behaviour*. Berlin: Mouton de Gruyter.

Seregély, Eva Maria. 2008. *A comparison of lexical learning in CLIL and traditional EFL classrooms*. Unpublished MA thesis, University of Vienna. [full text available at <[http://othes.univie.ac.at/2057/1/2008-10-13\\_0200428.pdf](http://othes.univie.ac.at/2057/1/2008-10-13_0200428.pdf)>]

Singleton, David. 1999. *Exploring the second language mental lexicon*. Cambridge: Cambridge University Press.

Sylvén, Liss Kerstin. 2004. *Teaching in English or English teaching?: on the effects of content and language integrated learning on Swedish learners' incidental vocabulary acquisition*. Unpublished PhD thesis. Göteborg: Department of English, Göteborg University.

Widdowson, Henry George. 1996. *Linguistics*. Oxford: Oxford University Press.

How to contact us:



c/o

**Institut für Anglistik & Amerikanistik der Universität Wien  
Universitätscampus AAKH, Spitalgasse 2-4, Hof 8 .3  
A – 1090 Vienna; Austria**

**fax** (intern.) 43 1 4277 9424  
**eMail** [views.anglistik@univie.ac.at](mailto:views.anglistik@univie.ac.at)  
**W3** <http://anglistik.univie.ac.at/views/>  
**(all issues available online)**

**IMPRESSUM:**

**EIGENTÜMER, HERAUSGEBER & VERLEGER:** VIEWS, c/o INSTITUT FÜR ANGLISTIK & AMERIKANISTIK DER UNIVERSITÄT WIEN, UNIVERSITÄTSCAMPUS AAKH, SPITALGASSE 2, A - 1090 WIEN, AUSTRIA. **FÜR DEN INHALT VERANTWORTLICH:** LOTTE SOMMERER, BARBARA SOUKUP **WEBMASTER:** MONIKA WITTMANN **REDAKTION:** HEIKE BÖHRINGER, ANGELIKA BREITENEDER, CHRISTIANE DALTON-PUFFER, OLGA FISCHER, CORNELIA HÜLMBAUER, THERESA-SUSANNA ILLÉS, GUNTHER KALTENBÖCK, KAMIL KAŻMIERSKI, URSULA LUTZKY, BARBARA MEHLMAUER-LARCHER, RUTH OSIMK, MARIE-LUISE PITZL, NIKOLAUS RITT, CLAUDIO SCHEKULIN, HERBERT SCHENDL, BARBARA SCHIFTNER, BARBARA SEIDLHOFER, EVA MARIA SEREGÉLY, UTE SMIT, LOTTE SOMMERER, BARBARA SOUKUP, JOHANN UNGER, H.G. WIDDOWSON, ANITA WOLFARTSBERGER. ALLE: c/o INSTITUT FÜR ANGLISTIK & AMERIKANISTIK DER UNIVERSITÄT WIEN, UNIVERSITÄTSCAMPUS AAKH, SPITALGASSE 2, A - 1090 WIEN. **HERSTELLUNG:** VIEWS