On abduction in receptive multilingualism. Evidence from cognate guessing tasks.

Raphael Berthele, University of Fribourg

This article focuses on the question how multilinguals use their languages in order to guess the meaning of cognates in unlearnt but genealogically close languages. A series of studies is discussed whose aim is to tap into this process of interlingual inferencing. Different measures for phonological and graphematic distances across languages are established and correlated with the rates of successful cognate recognition in the search for a threshold of string similarity beyond which recognition becomes unlikely. The role of different types of the participants' multilingual repertoires is assessed, and other factors influencing good performance in cognate recognition are identified. The process of interlingual inferencing is discussed as a form of abductive reasoning, and quantitative and qualitative data are analyzed to support the idea that this type of abduction is an essential driving force in receptive multilingualism and language comprehension in general.

1. Inferencing and abductive reasoning in language learning and use

There is converging evidence for the claim that bi- and multilinguals have, at least in some respects, advantages over monolinguals in learning additional languages (Cummins 2000: 35; Cenoz 2003: 82). The explanations offered for these advantages in third language acquisition (TLA) refer to advantages in cognitive development or to enhanced language awareness (Jessner 1999), sometimes also termed the "M-factor" (Herdina & Jessner 2002: 131): Learning and using two languages is hypothesized to lead to enhancements in particular cognitive skills and/or language awareness, which in turn leads to measurable advantages – at least regarding general efficiency – in TLA (cf. also de Angelis 2007; Le Pichon Vorstman, De Swart, Ceginskas & Van Den Bergh 2009).

The empirical studies presented in this chapter focus on a narrowly defined aspect of linguistic competence, the recognition of cognates in closely related but unknown languages. More concretely, this research attempts to tap into a cognitive mechanism which is likely to contribute to the advantages in TLA within genealogically related languages observed by

scholarly research in the past. The goal of the deliberately reductionist approach discussed

below is to shed new light on a pivotal resource in multiple language learning: Interlingual

correspondences. The research questions that will be addressed in this chapter are the

following:

1) Are there multilingual profiles (regarding the languages in the individual multilingual repertoire) that favor the rapid recognition of familiar words in unknown target languages?

2) Are there other characteristics of the participants that influence the quality of the interlingual recognition process?

3) Which linguistic properties of target items (words) allow for spontaneous interlingual recognition, and which limit the spontaneous inferability of items?

These three questions will be addressed using data elicitation methods described in Section 2.

Section 1 provides a brief overview of the context of this research and the most relevant

literature, and Section 3 offers some concluding remarks.

The main point of this chapter is theoretical. It will be argued that one of the key inferencing

procedures that underlie the linguistic tasks in the scope of our studies is of the abductive

type. The particular nature of abduction, then, can be used as an explanation as to why bi-

/multilinguals do better at the particular tasks we used and, as a consequence of this first

conclusion, why many bi-/multilinguals they might be better at language learning in general.

1.1 *M-Factor; bi-/plurilingual competence*

A focus on cross-linguistic similarities is becoming increasingly important as a means of

promoting the usage value of less commonly used languages (LOTE, languages other than

English) in light of an over-emphasis on teaching English. As one of many examples, the

EuroCom framework (McCann, Klein & Stegmann 2003; Hufeisen & Marx 2007) aims at the

rapid development of comprehension skills in reading and listening in one family of

languages at a time (Romance or Germanic). Picking up on an idea already present in Lado's

(1957) contrastive analysis hypothesis, these frameworks pursue a double focus: On the one

hand, the potential of interlingually transparent words and structures is documented, on the

other, particular features that are unique for the particular language are listed and presented.

Although the general idea of this enterprise seems very promising, one of its main weaknesses

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.
p. 3

is its orientation towards a detailed, lengthy, list-based presentation of interlingual correspondences that draws on the philological tradition. The usability of these lists remains at best unclear. As an example, sound correspondences across Germanic languages are presented in EuroComGerm (Hufeisen & Marx 2007) on no fewer than 44 pages. Given the density of these contents, and despite the fact that the authors have made great efforts to reduce the philological complexity, it seems that there should be better, more frugal ways of preparing multilinguals for the task of reading in an unknown but genealogically familiar language.

1.2 *Inferencing, transfer and hypothesis testing*

According to Carton (1971), one of the fundamental principles in language learning is the ability to draw inferences. Studer (2008) deems inferencing a central process in almost any receptive activity of language users. Inferences can be drawn based on co-textual or contextual information, based on knowledge of the target language, of languages in general or based on knowledge that links different languages in the multilingual system together. It is these latter inferences, the interlingual ones, which are at the very core of this chapter. Furthermore, such interlingual inferences are closely related to central questions on the potential (but also the danger) of language transfer, as discussed e.g. in Odlin (1989). For the remainder of this contribution, we consider interlingual inferencing as a mechanism that operates on potential transfer bases, i.e. on acquired and learnt items and structures in any language pertaining to the multilingual repertoire. Based on inferentially emerging interlingual identification, multilinguals can be said to transfer items and structures from one language into another. Both inferencing and transfer therefore participate in the process of the dynamic construction of interlanguages in the form of often short-lived, spontaneous, hypothetical grammars, a process that is considered central to most modern theories of second/foreign language acquisition (see e.g. Gass & Selinker 1994: 6-7) as well as in multilingual language didactics (see e.g. Meißner 2001). However, the exact nature of these inferencing processes often remains surprisingly unclear. Most often, particularly in so called 'constructivist' teaching approaches, the emphasis lies on discovery, on 'inductive' procedures. These procedures are argued to be more efficient and more cognitively adequate for language learning than the traditional deductive or 'instructivist' methods (cf. Wolff 1994).[1] In this regard, however, authors tend to forget that there is a third type of inferencing, namely

---

[1] Radically inductive approaches to foreign language teaching have been criticized for quite some time now, and the pendulum is swinging back to approaches that include focus on form or to mixed 'post-method era' practices. However, especially in the German-speaking world, the idea of radically 'constructivist' foreign language teaching is still surprisingly alive and kicking.

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.
p. 4

abduction.[2] Abduction was introduced into semiotic theory by Peirce (1931) who was interested in the "mental operation of guessing" and viewed abduction as a process of "forming an explanatory hypothesis" (CP 5.171). In the view of Eco (1984), abduction is the daring attempt to speculate about potential rules that might explain the meaning of a signifier (42).
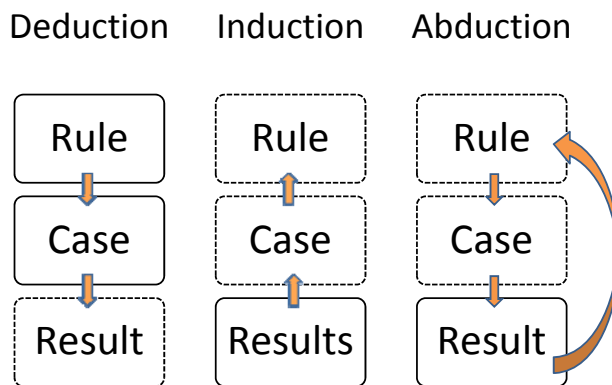

Figure 1: Three types of inferencing according to Eco (1984: 42). Bold lines refer to given knowledge, dotted lines refer to inferenced knowledge.

Let us consider a token from the data collected in the studies discussed below and how it relates to the three types. One of the tasks involved the spontaneous guessing of the meaning of cognate verbs in Danish, a language none of the informants has learnt. As an example (1), the participants had to infer the meaning of the item *blive*, an item that was presented either aurally or in its written form (but not both).

(1a) Aural-oral variant: [bli:ve] – What does this Danish verb mean?

Participant: From French *plier* [plie:] or from English *believe*, [pliɐ]

Interviewer: What's that in English?
Participant: 'believe', but this doesn't fit in
Interviewer: Why? […]
(1b) Written task: What does the Danish verb <blive> mean?
bleiben; in Spanish v and b are often pronounced the same, […] thus „BLIBE"→ German *bleiben* ('to stay')
Both examples are actual responses drawn from the study that will be described in more detail below. Example 1b is a vivid example of interlingual abduction: The participant is confronted

---

[2] Abduction has been mentioned before in linguistics: firstly as a mechanism in language change (Mc Mahon 1994), and secondly by Chomsky (1968) as a principle in first language acquisition.

with a stimulus, she has a gut feeling that it could be related to the German *bleiben*, and she looks for potential phonological 'rules' that might support and even explain this spontaneous analysis, using her knowledge of Spanish grapheme-phoneme correspondences. Based on Eco's schematic representation of induction, deduction and abduction, this particular token can be rephrased as in Figure 2.

Deduction   Induction   Abduction

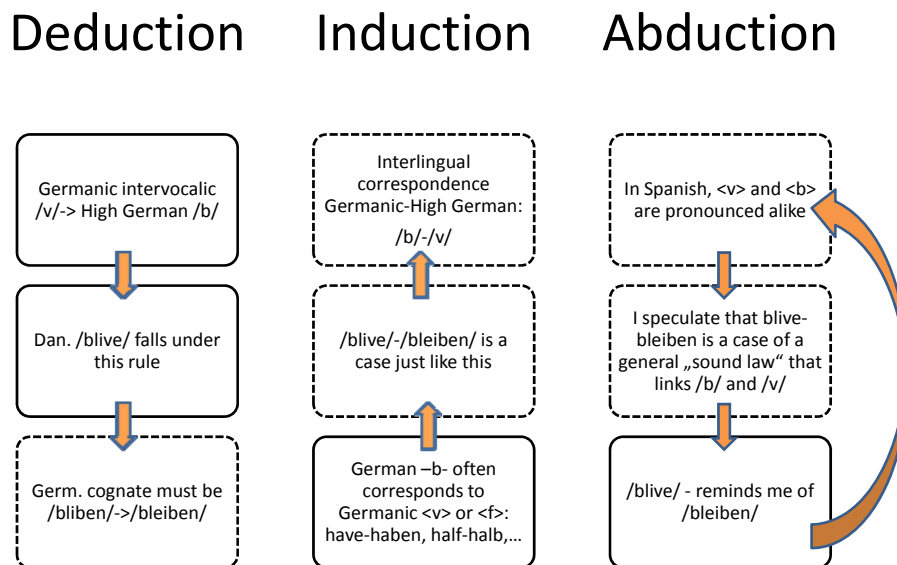| Deduction	Induction	Abduction
Germanic intervocalic /v/-> High German /b/	Interlingual correspondence Germanic-High German: /b/-/v/	In Spanish, <v> and <b> are pronounced alike
Dan. /blive/ falls under this rule	/blive/-/bleiben/ is a case just like this	I speculate that blive-bleiben is a case of a general „sound law" that links /b/ and /v/
Germ. cognate must be /bliben/->/bleiben/	German –b- often corresponds to Germanic <v> or <f>: have-haben, half-halb,...	/blive/ - reminds me of /bleiben/

Figure 2: Deduction, induction and abduction exemplified with Germanic sound correspondences.

As Figure 2 shows, there are other ways of linking together the rule, the case and the result. In deduction, the result is an unavoidable entailment of the rule and the case, and in induction results can allow for the inferential generalization of a rule. The figure should make it quite clear that both induction and abduction by nature involve uncertainty regarding the validity of the inference as a whole to at least some degree. This uncertainty is particularly important in abduction, since here the empirical basis is rather weak (often only 1 result). A second difference between abduction and induction needs to be explicitly mentioned here: The former presupposes the knowledge of rules, as the one given in example (1b), whereas the latter is about finding the rule based on a corpus of data. This difference, as will be argued in this article, is a crucial one – from an acquisitional and from a teaching point of view.

2. Empirical Investigations

2.1 *Lexical inferencing across Germanic and Romance languages*

Two preliminary studies were carried out that will be summarized very briefly here (see Berthele 2008; Berthele & Lambelet 2009 for more details). Both studies were inspired by the methodology developed in Müller-Lancé (2003). The studies make use of paper and pencil tasks targeting unlearnt languages that are genealogically related to the main languages of the participants. More importantly, the target languages and the learnt languages share a great deal of vocabulary. These interlingually shared words, commonly referred to as 'cognates', differ to substantial degrees with respect to their similarity. It seems thus important to deconstruct the lay category of cognates and to reconstruct it based on scientific criteria (cf. Section 2.4).

	Study 1a	Study 1b
N (# of participants)	183	150
sample	Students of German at the Universities of Zürich, Marburg, Fribourg/Freiburg	Students of Psychology at the University of Fribourg/Freiburg
participants' L1	German (dialect or standard)	French or Italian
mean # of languages in the repertoires	4.25	4.29
Task A: Infer meaning of words (Word list with context)	17 Dutch words (from a reading text, authentic article, cognates and non-cognates)	17 Romontsch Sursilvan words (from a reading text, translated newspaper article, cognates and non-cognates)
Task B: Reading comprehension	Dutch reading comprehension (7 questions)	Romontsch Sursilvan reading comprehension (7 questions)
Task C Infer meaning of cognates (Word list without context)	29 Danish/Swedish verbs, all high frequency, all (near) cognates	29 Romanian and Romontsch Sursilvan verbs, all high frequency, all (near) cognates

Table 1: Two structurally analogous studies on Romance and Germanic target varieties

All participants had to fill in a detailed language biography questionnaire in order to self-assess their proficiency in their languages (including dialects). The tasks were carried out

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.

under time pressure (Task A: 15', Task B: 5', Task C: 10'). The translation attempt for each item in an unknown Lx involved writing down the (guessed) meaning of the item as well as the interlingual (or contextual) transfer bases, words or rules, which lead to the guess (cf. example 1b).

2.2   Results Studies 1a and 1b

Firstly, a general correlation between the number of languages spoken by the participants and the number of successful translation attempts was calculated. This analysis reveals a positive but relatively weak effect for task C only (inferencing of cognates without context; Romance targets: $r = 0.392$, $n = 135$, $p < 0.001$; Germanic targets: $r = 0.249$, $n = 179$, $p<0.001$). Secondly, the question was addressed as to whether there are particular multilingual profiles that stick out with respect to the ability to draw interlingual inferences. Different groups were formed, e.g. a group of bilinguals has been defined as participants who speak at least two languages/dialects on a level of at least 4 (maximum: 6; labeled "(almost) balanced bilingual" in Figure 3). Another group of 'intra-family' bilinguals was defined as speaking at least two Germanic (study 1a, "2 Germanic >4" in Figure 3) or Romance (study 1b, "2 Romance >4 in Figure 3) languages on a level of at least 4. Individuals not falling into these categories are labeled "(normal) multilingual" in Figure 3. There is considerable evidence on both the Romance and the Germanic side that the most skillful interlingual inferers are those participants who master two languages in their repertoire that are related to the target languages, i.e. two Romance or two Germanic languages respectively (cf. Figure 3). On the Germanic side this shows the multilingual potential that lies in bilingualism with dialects and a standard language, since a large number of the participants who claim to have high proficiency in two Germanic systems are speakers of a German dialect (Alemannic, Platt, etc.) and the German standard language (cf. the discussion in Berthele 2008).

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter. p. 8
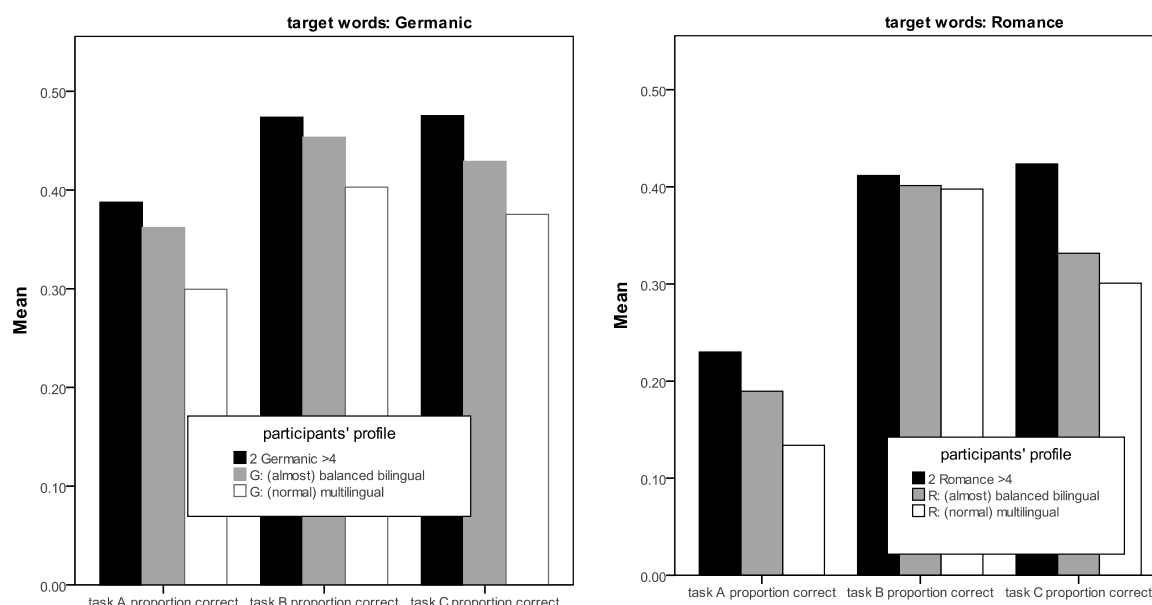
Figure 3: Multilinguals' profiles and respective success in tasks A, B, and C.

To sum up the result of the search for the "good interlingual inferer" we can conclude that:

a) the more multilingual a participant is, the more likely she is to correctly infer word meaning (all three tasks in the case of the Germanic targets, tasks A and C in the case of the Romance targets);

b) multilinguals who assess themselves as having a high proficiency in at least two of their languages do better than multilinguals without this particular type of self-assessment;

c) multilinguals with high proficiency in two languages that are close to the target languages perform better than all other groups.

In order to be able to increase the amount of variance explained, we designed a follow-up study described in the remainder of this chapter. This study draws on the method used in task C.

2.3   *Study 2: Listening and reading comprehension in Germanic languages*

The follow-up study only focuses on Germanic verb targets without context. A total of 163 adolescents and young adults (between 13 and 35 years old) participated, all native speakers of a Swiss German dialect (or bilinguals with heritage languages or Romansh). They had to fill in a similar language profile questionnaire as in study 1. Additionally, the informants took 3 modules of Meara's (2005) language aptitude tests (called *Llama B, D* and *E*: word learning,

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter. p. 9

sound recognition, sound-symbol correspondence). The stimuli were 28 Danish and Swedish verbs, all high frequency and (near) cognates, i.e. they are potentially detectable with counterparts in English and/or German (cf. below, Section 2.4).

The informants were presented with half of the stimuli in written form and half in aural form. Two conditions were created in order to have all verbs presented in both forms. A subsample of the informants performed the same task in a think-aloud protocol, i.e. not as a paper and pencil task but face to face with a fieldworker who recorded the verbal protocols and wrote down the responses. Example (1a) above is taken from this data.

The data have been analyzed in two ways: firstly with respect to the quest for the profile of the 'ideal' interlingual inferer, and secondly with respect to linguistic features that interact in statistically relevant ways with the empirical difficulty of the stimulus items.

In order to answer the first question, a regression analysis was carried out. The dependent variable was the standardized score of the proportion of correct inferences an individual managed to draw. As independent variables, supposedly important factors were entered stepwise: vocabulary learning ability (Llama B), sound recognition (Llama D), sound-symbol correspondence (Llama E), number of languages in the repertoire, self-assessed proficiency in L1, L2, …, Lx, age, school/educational level.

The stepwise regression reveals that there are four independent variables that contribute in a statistically meaningful way to the variance of the target variable (cf. Table 2):

1. age (the older, the better), 2. vocabulary learning ability, 3. English proficiency, 4. # of languages in the repertoire. These four variables account for 62% of the variance.

adjusted R^2 = .615 Model :	B	Std. Error	Beta	R^2 change	Sig.
1 Age	.196	.044	.587	.344	.000
2 Age	.167	.041	.501		.000
Llama B - vocabulary learning	.018	.006	.393	.147	.003
3 Age	.110	.045	.330		.019
Llama B - vocabulary learning	.019	.005	.420		.001
English proficiency	.280	.116	.318	.073	.021
4 Age	.089	.044	.267		.050
Llama B - vocabulary learning	.016	.005	.356		.004
English proficiency	.310	.111	.352		.009

# of languages in the repertoire	[3]-.236	.111	-.244	.051	.042

Table 2: Stepwise regression coefficients for the number of successful inferencing attempts

The analysis shows that the ability to draw inferences increases with age, that there is at least one component of language aptitude that seems to interact with this ability, and that the multilingual repertoire, and most prominently the proficiency in English, contributes significantly to successful inferencing.

If we consider these results in the light of the presupposition that we are dealing with abduction rather than induction, the results make perfect sense: Multilinguals with high proficiency in many languages, and above all in languages that are relatively similar to the target language, can make use of their repertoire in the sense that it provides not only important lexical transfer bases, but also potential rules and regularities, i.e. conceptual knowledge and strategic know-how that fits into the top box in Figure 1. It is this knowledge and these skills that allow the interlingual guessing game to become increasingly accurate, as we will see in much more detail in Sections 2.5 and 2.6.

Before discussing these results and their potential consequences for the acquisition and learning of third/additional languages, another question should be briefly addressed: What is the influence of particular linguistic characteristics in the target words on successful interlingual comprehension?

2.4   *What exactly is a cognate? Analysis of individual Items*

The term *cognate* is well established in psycholinguistic investigations of the multilingual mental lexicon (e.g. in Dijkstra 2003), and the recognition of interlingual cognates is taken for granted and thus part of an implicit methodology in comparative historical linguistics (for an example see Campbell 2004: 126).

---

[3] It is only after the paper went into print that I realized that the discussion of the 'multilingualism factor' is not sufficiently explicit, since it does not address the question why this coefficient is negative: More languages in the repertoire correlate with a smaller amount of correct inferences. Obviously this result goes against my expectations, and at this point in time I have no idea what kind of theory would predict this kind of result.

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter. p. 11

Despite its being taken for granted, *cognate* is a rather problematic category. There are clear cases of obvious code similarity (homography or homophony) across languages, such as the Danish and the English spelling of the verb <give>. In other cases, however, minimal or even major differences in pronunciation occur, and where 'cognateness' ends often remains unquestioned in the literature. Both in spelling and in pronunciation, even within what is commonly construed as a language, there can be substantial allophonic and allographic variation. I would suggest conceiving the category of cognate as a radial category with prototypical, i.e. almost or completely identical examples at its center and increasingly different examples on a graded scale. Moreover, the category has fuzzy boundaries since it remains unclear where interlingual cognateness ends. As shown in Figure 2 above, knowledge of the history of sound changes in Germanic and synchronic sound correspondences allow for the identification of graphematically and phonologically quite different items.

Figure 4: The radial category of synchronic cognateness

The construal of cognateness as a radial category avoids the pitfall of assuming the interlingual indentifiability of items in an aprioristic way. In order to account for the gradedness of the cognate category, the interlingual distances need to be measured. Once these measures are established, it is possible to correlate them with the empirically measured ease of inferencing in our experiments.

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.
p. 12

Measuring string distance is possible by applying a variant of the algorithm proposed by Levenshtein (1966). Whereas this algorithm, in its simplest form, only counts the smallest number of insertions or deletions that have to be carried out in order to transform one string into the other, there are elaborations that take also into account the phonological distance between the respective segments (Heeringa, Kleiweg, Gooskens & Nerbonne 2006). For the purposes of this study, I have used two variants of the Levenshtein algorithm. Firstly, a featureless comparison of the graphematic strings has been carried out (cf. columns 2 and 4 in Table 3). This produces a rough measure for the graphematic similarity of the two written forms of the cognates. Secondly, the aural stimuli of the words have been transcribed phonetically, these transcripts have been coded using the X-Sampa[4] standard and fed into the L04[5] software, which is a tool that calculates feature-based Levenshtein distances for word pairs.

	kunne (Danish)		*mena* (Swedish)	
	Graphematic level	Phonological level	Graphematic level	Phonological level
Target	<kunne>	[ˈkʰunə]	<mena>	[ˈmiɛˈnɑ]
Cognate German	<können>	[ˈkʰoenːən]	<meinen>	[ˈmaɪnən]
Cognate English	<can>	[kʰæːn]	<mean>	[miːn]
English: distances	0.80[6]	0.24	0.40	0.29
German: distances	0.33	0.14	0.50	0.14
Empirical difficulty	0.453	0.161	0.557	0.081

Table 3: Example items with Levenshtein distances

Very generally, it seems reasonable to hypothesize that the items' empirical difficulty in the interlingual inferencing task correlates negatively with the Levenshtein distances. In order to test this hypothesis, the proportion of correct answers per item per task variant has been calculated (cf. last row in Table 3). A correlation analysis of these values with the Levenshtein distances shows a negative interaction with the distances between the target item and the English cognate for both the written and the audio stimuli (written: r = -0.416, n = 28, p=0.014; aural: -0.349, n = 25, p = 0.026). No such interaction can be found correlating the

---

[4] cf. http://www.phon.ucl.ac.uk/home/sampa/x-sampa.htm [July 28, 2010]

[5] cf. http://www.let.rug.nl/~kleiweg/L04/ [July 28, 2010]

[6] The feature-based and featureless distances have both been normalized using the longest length among the set of least cost alignment solutions, as recommended in Beijering et al. (2008).

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.

p. 13

target and the German values (written: r = 0.14, n = 28, p=0.471; aural: r = -0.002, n = 28, p = 0.497).

2.5 *Listening comprehension: Analysis of individual Items*

The data from the listening comprehension condition will first be discussed.

As Figure 5 shows, there seems to be a threshold at around 0.22 for the distance measure in the listening comprehension condition. Above this threshold, no items were inferred by more than 15% of the participants.

Figure 5: Inferability of verbs and Levenshtein distances to English

Levenshtein distances with respect to English thus seem to be a modestly useful predictor for the inferability of the target items. However, as Figure 5 illustrates, there is no simply linear relationship between the two variables. Among the items with relatively small phonological distances we still find considerable variation. Some items are comparatively distant but still easy to infer, e.g. *komma*. In some case, this can be interpreted due to the high similarity to German (*kommen*), in others, however, this explanation does not seem to be equally manifest

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.
p. 14

(*ville – wollen*). The lack of clear correlations and the unexplained patterns in Figure 5 raise the question whether it is not mere distance/difference, but rather particular types of differences that are the key to the empirical differences across the items in this interlingual guessing puzzle. A further explorative analysis was thus carried out that analyzes phonological and graphematic features of particular segments of the pairs of cognates. Firstly, since the data points in Figure 5 suggest that there is a meaningful difference with respect to interlingual inferencing between items with a Levensthein distance of more than 0.218, the target items have been categorized in two groups (low Levenshtein distance D<0.218; high Levenshtein distance D>0.218). Secondly, a more elaborate item-map has been constructed for each item in the list. The simplified schematic structure of the target items is C-V-C-V pattern (cf. *kunne*, *mena*). In some cases, the consonant positions are filled with consonant clusters, as in *skrive*. For each of these four positions, phonetic features have been coded based on the phonetic transcriptions of the items. The item-map thus contains variables which code the main articulatory differences between the cognates for each pair of words:

Consonants: place/manner of articulation; phonation
Vowels: aperture; backness; roundedness; diphthong/monophthong;
Vowels and Consonants: segment differences (e.g. affricates vs. fricatives; insertions or deletions of sounds); quantity

If there is more than one segment per idealized C or V position, as in the case of consonant clusters, the differences between these clusters have been coded as a whole. Based on this item-map with an impressive number of variables (54 for the aural data) it is possible to run a cross-tabulation analysis which indicates whether each particular feature coincides with higher or lower success in interlingual inferencing. An exploratory analysis has been carried out using classification trees with the dependent variable "correct inference" (a nominal variable) and the whole phonological item-map as independent variables. The goal was to find out which of these numerous variables discriminate best with respect to the dependent variable.[7] In other words, the question here is which linguistic feature clearly brings about less

---

[7] Additionally, this method acts as a for the total number of alpha errors when running multiple statistical tests. For nominal dependent variables, the QUEST algorithm is the recommended procedure.

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter. p. 15

success in inferencing when two cognates differ with respect to it. This analysis was carried out for two groups of participants separately: for the good inferers and for the poor inferers. These groups were determined by dividing the sample according to the criterion of whether a participant scored below or above the mean of correct inferences (36%) for the whole sample. This analysis produces the picture shown in Figure 6 and Figure 7.
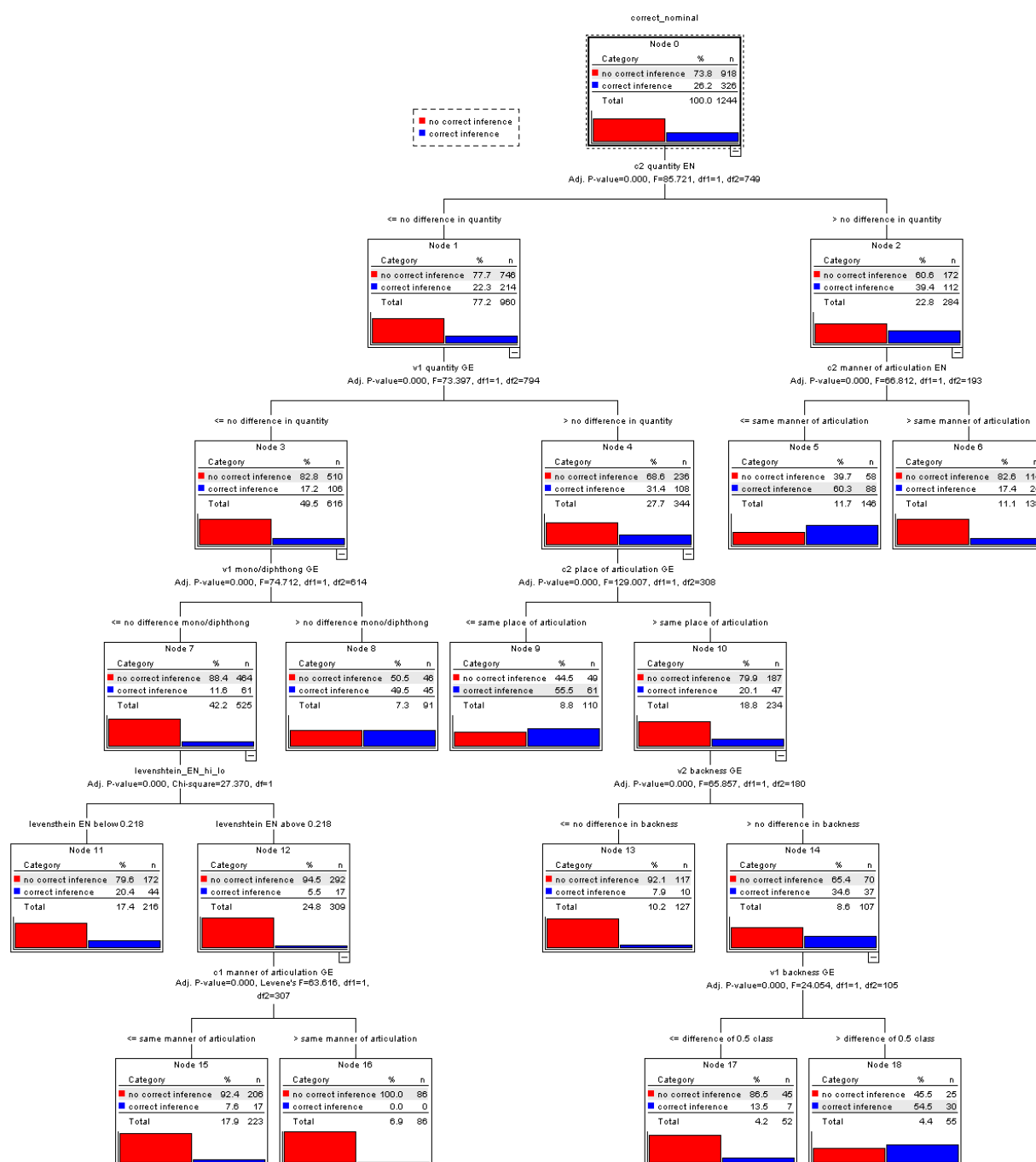

Figure 6: Listening comprehension (bad inferers)

correct_nominal

Node 0

Category	%	n
■ no correct inference	56.6	554
■ correct inference	43.4	424
Total	100.0	978

□ no correct inference
□ correct inference

levenshtein_EN_hi_lo
Adj. P-value=0.000, Chi-square=132.347, df=1

levensthein EN below 0.218

Node 1

Category	%	n
■ no correct inference	43.5	279
■ correct inference	56.5	363
Total	65.6	642

levenshtein EN above 0.218

Node 2

Category	%	n
■ no correct inference	81.8	275
■ correct inference	18.2	61
Total	34.4	336

c2 phonation EN
Adj. P-value=0.000, F=104.367, df1=1, df2=348

<= no differnce in phonation

Node 3

Category	%	n
■ no correct inference	37.3	164
■ correct inference	62.7	276
Total	45.0	440

> no differnce in phonation

Node 4

Category	%	n
■ no correct inference	56.9	115
■ correct inference	43.1	87
Total	20.7	202

v1 aperture GE
Adj. P-value=0.000, F=52.821, df1=1, df2=404

<= difference of 1step

Node 5

Category	%	n
■ no correct inference	22.6	52
■ correct inference	77.4	178
Total	23.5	230

> difference of 1step

Node 6

Category	%	n
■ no correct inference	53.3	112
■ correct inference	46.7	98
Total	21.5	210

v1 roundedness EN
Adj. P-value=0.000, F=89.419, df1=1, df2=200

<= no difference in roundedness

Node 7

Category	%	n
■ no correct inference	85.6	83
■ correct inference	14.4	14
Total	9.9	97

> no difference in roundedness

Node 8

Category	%	n
■ no correct inference	30.5	32
■ correct inference	69.5	73
Total	10.7	105

c1 manner of articulation EN
Adj. P-value=0.000, F=39.496, df1=1, df2=228

<= same manner of articulation

Node 9

Category	%	n
■ no correct inference	11.5	18
■ correct inference	88.5	138
Total	16.0	156

> same manner of articulation

Node 10

Category	%	n
■ no correct inference	45.9	34
■ correct inference	54.1	40
Total	7.6	74

c2 different # of segments EN
Adj. P-value=0.000, F=114.129, df1=1, df2=208

<= same number of segments

Node 11

Category	%	n
■ no correct inference	31.9	44
■ correct inference	68.1	94
Total	14.1	138

> same number of segments

Node 12

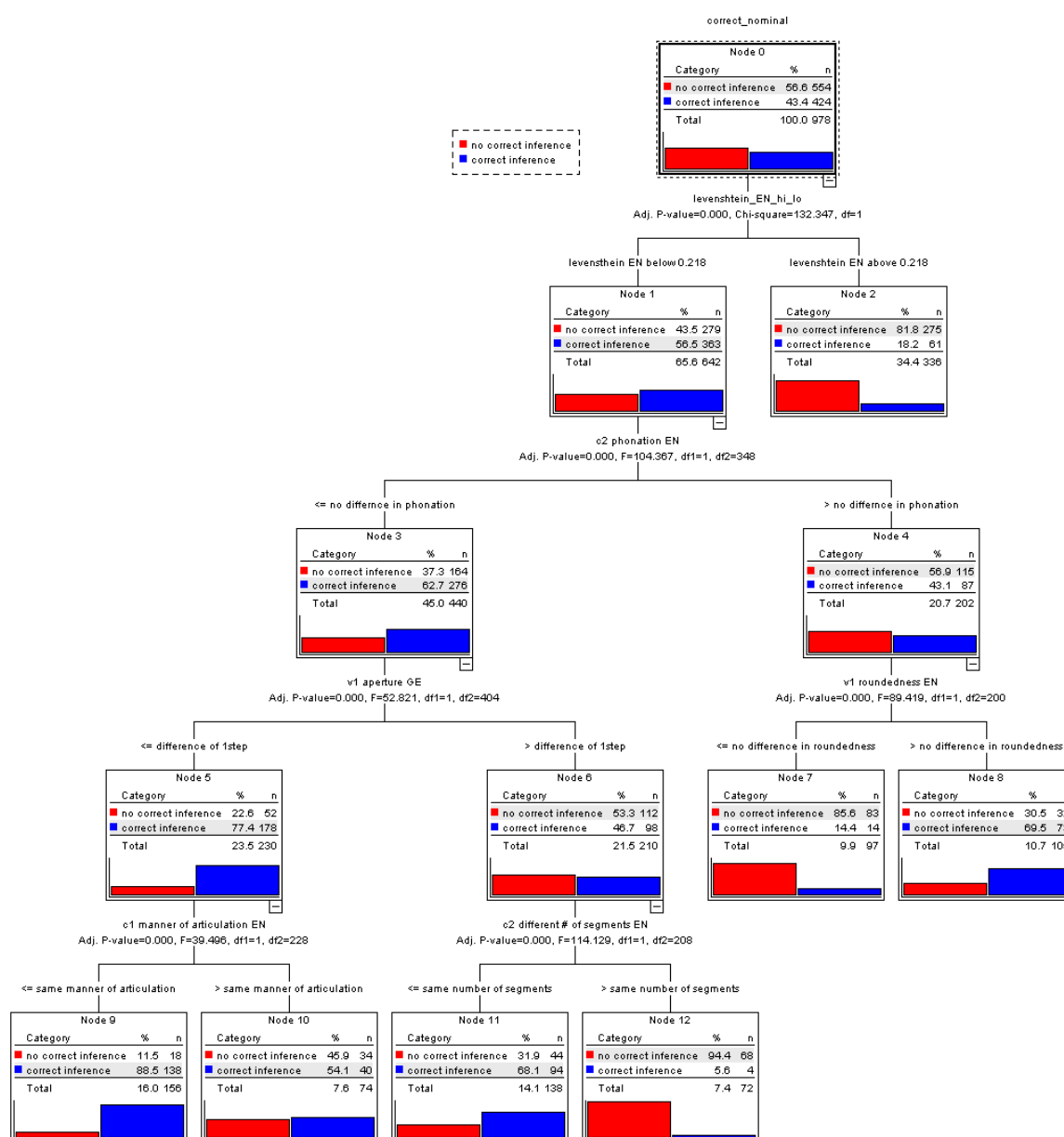| Category	%	n
■ no correct inference	94.4	68
■ correct inference	5.6	4
Total	7.4	72

Figure 7: Listening comprehension (good inferers)

The exploration of these categorization trees shows which phonological features discriminate within our data with respect to successful interlingual inferencing.

In order to lay out the logic of these trees, two selected paths that lead through the tree in Figure 7 will be analyzed as an example.

Firstly and most importantly, if an item has a Levenshtein distance that is higher than the threshold discussed above, it is very unlikely for the good inferers to find the cognate. Only 61 (18.2 %) of the items are successfully translated, 275 (81.8%) remain unrecognized. However, if the Levenshtein distance is below the threshold, the good inferers succeed in 363

attempts (56.5%). The second criterion that plays a statistically significant role is the comparison of phonation of the second consonant in English and in the target item. If there is a difference in phonation, or if one of the forms lacks a second consonant, it is less likely that the inferences will be correct (43.1% of correct inferences). Within these cases, the first vowel criterion is important: If the first vowel shows a difference in roundedness when comparing English and the target item, the inferences are more likely to be successful than if not. Node 8 thus stands for verbs such as ha [hɒː], gå [ˈɡoːə], and stå [ˈstɔʔ], that are relatively well translated (69.5%).[8]

Let us now turn to the items that do not present any difference in phonation regarding the second consonant, i.e. to all items that belong to daughters of node 3. The next criterion that plays an important role is the comparison of the aperture of the first consonant in the German and the target cognates. If there is no difference at all or merely a minor one (node 5), 77.4% (178 tokens) are correctly inferred. If there is a difference larger than 1 step, the majority of the items is not inferred correctly (53.3%). Analyzing the daughters of the former category (node 5 in Figure 7), the last criterion is the manner of articulation comparing the first English consonant and the target words. If the words present the same manner of articulation, there is an 88.5% likelihood that they will be inferred correctly (items such as komma, *lægge, ligge*). Even if this feature differs (node 10), the majority of the tokens are correctly inferred (*tale, tänka*; 54.1% correct inferences).

All nodes in the two figures above could be analyzed in this manner. For the sake of brevity we will not go into more details here but the following table shows a synopsis of the hierarchy of the features that the categorization tree analysis produces as being statistically significant for the good and bad interlingual inferers.

	good inferers				poor inferers			
rank of constraint	consonant vowel Levenshtein	EN GE	Feature	expected or unexpected relationship	Consonant Vowel Levenshtein	EN GE	Feature	expected or unexpected relationship
1	L	EN		expected	C2	EN	quantity	expected

---

[8] These examples illustrate an important flaw in the method applied, a difficulty that at this point has not been satisfactorily resolved: All of the participants are not only proficient in standard German, but also in one or several Alemannic Swiss German dialects as their respective L1s, many of which have cognate forms that are quite close to the target items (e.g. Zurich German /hɒː/ 'have', /ʃtɒː/ 'stand'). However, since the different Alemannic dialects spoken by the different informants are quite different in phonology, it turned out to be impossible to construct an analogous item-map for the dialectal cognates. The study discussed here thus, in future variants, should control systematically for the native dialects and include them in the analyses.

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.
p. 18

2a	C2	EN	phonation	expected	V1	EN	quantity	unexpected
2b					C2	EN	manner of articulation	expected
3a	V1	GE	aperture	expected	V1	GE	monophthong/ diphthong	unexpected
3b	V1	EN	roundedness	unexpected	C2	GE	place of articulation	expected
4a	C1	EN	manner of articulation	expected	L	EN		expected
4b	C2	EN	# of segments	expected	V2	GE	backness	unexpected
5a					C1	GE	manner of articulation	expected
5b					V1	GE	backness	unexpected

Table 4: Important structural characteristics in listening comprehension

Table 4 provides a hierarchically structured account of the features that interact in a statistically significant way with the probability that participants carry out correct inferences. The table can be used for a tentative exploration of structural constraints governing the interlingual inferencing task.

1. The most important factor for the good inferers is the Levenshtein criterion, i.e. whether an item lies below or above the threshold visible in Figure 7 above. Good inferers are thus relatively likely to infer items that are below this threshold but highly unlikely to infer items that lie beyond this threshold. This criterion only shows up in fourth position for the group of poor inferers. This means that these participants do a relatively poor job even if the Levenshtein distance is below this threshold.

2. Of the 6 features listed in Table 4 that relate to vowels, 4 have an unexpected relationship regarding the phonological difference between the cognates and the number of correct inferences; in other words, difference here coincides with more correct inferences.

3. All consonant differences between the cognates listed in Table 4 stand in the expected relationship with the number of correct inferences.

4. Of the 15 features listed in Table 4, 7 relate to consonants, 2 to the Levenshtein criterion, and 6 to vowels. One could tentatively hypothesize that consonants seem to play a slightly more important role, but only when taking into account the unexpected relationship between vowels and correct inferences discussed in (2) and the invariably expected effect direction between consonants and correct inferences (3).

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.
p. 19

5. Of the 15 features listed in Table 4, 9 relate to English and 6 to German. Moreover, the first two levels in the respective hierarchies are always occupied by feature comparisons to English.

6. Good inferers infer correctly if consonants are the same (3 C variables vs. 2 V variables, one of which behaves inversely). Poor inferers, at first glace, show a less clear picture - 4 variables are related to vowels, 4 to consonants. Delving deeper, however, it turns out that all four vowel variables show unexpected patterns, i.e. the vowel differences coincide with higher levels of correct inferences compared to word pairs with the same vowels. This suggests an even stronger role of consonants: If they are the same, bad inferers have a relatively high chance of inferring the words correctly, but differences in vowels do not seem to trigger worse performances.

2.6 *Reading comprehension: Analysis of individual Items*

Before commenting further on these exploratory results, let's briefly turn to the analogous analysis of the written task.

For the reading comprehension data, the coding was simply based on the question of whether the particular segment is same, similar (v-w, b-p, etc.) or different (t-s). Since we have little access to the participants' imagined phonology when reading the words, the vowels were strictly coded in a nominal and rather approximate way for sameness or difference, erasing all diacritics and length markers in all languages (i.e. German <ie> is treated as <i>, <ä> as <a>, Swedish <å> as <a>, etc.).

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter. p. 20
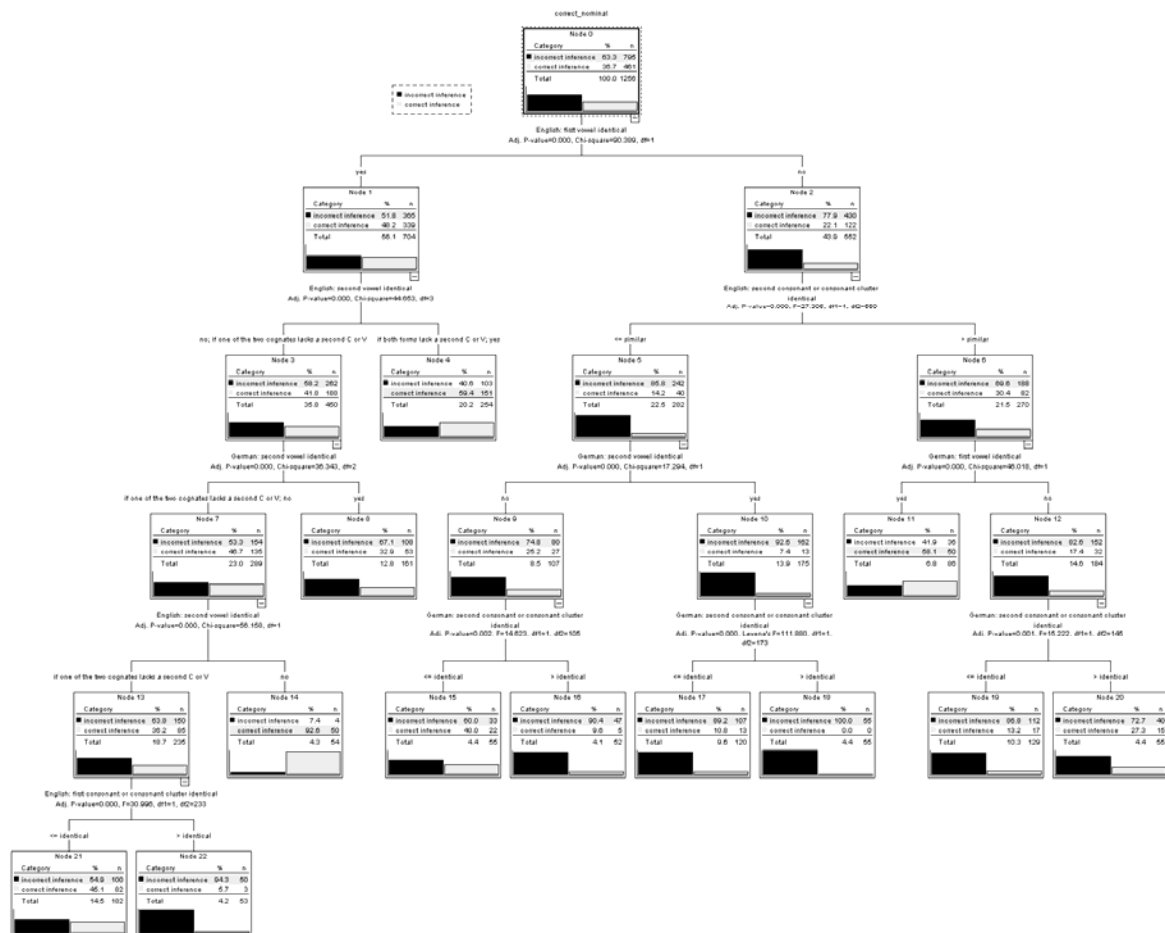

Figure 8: Reading comprehension (bad inferers)

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter. p. 21
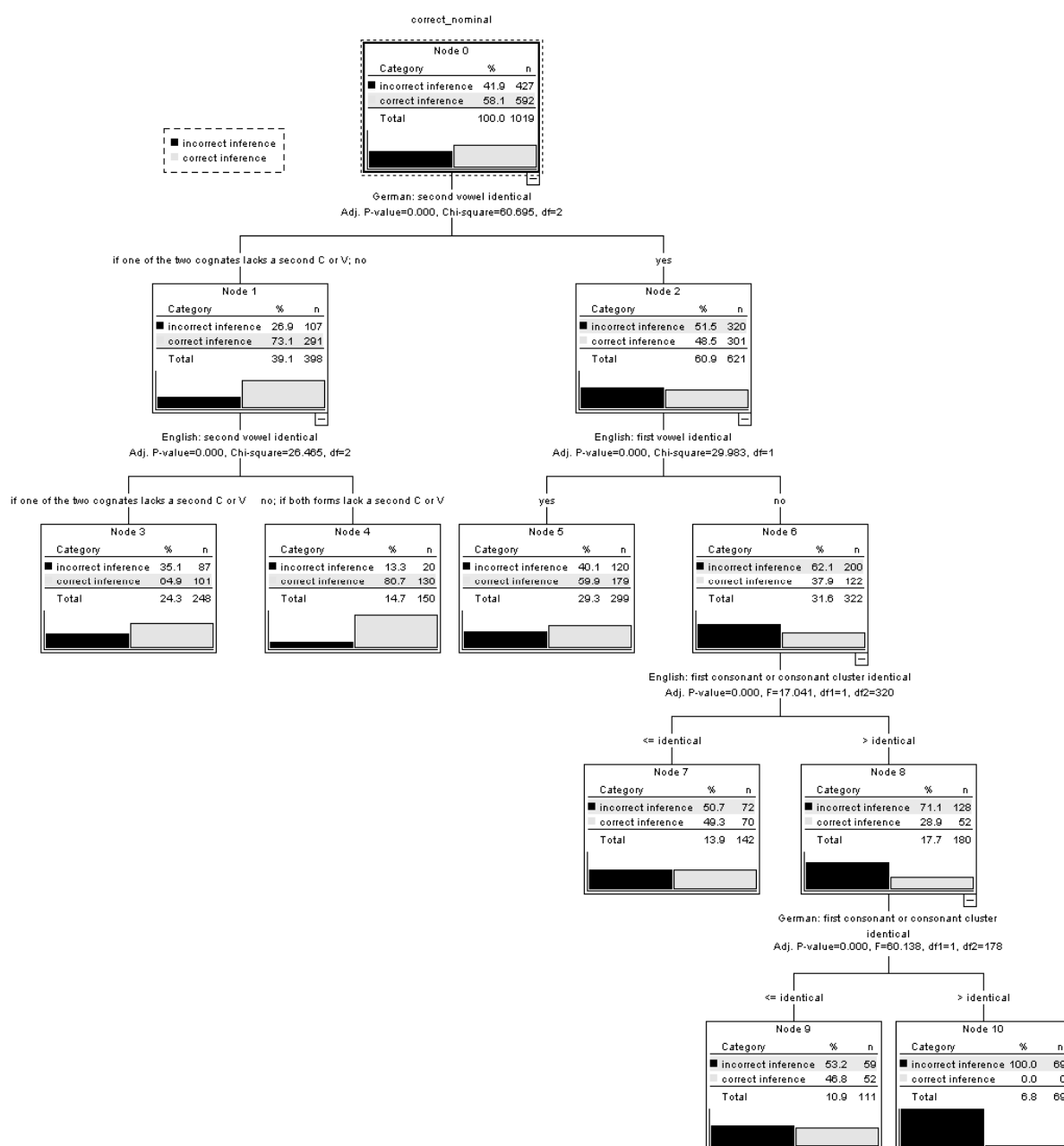
correct_nominal

Node 0

Category	%	n
■ incorrect inference	41.9	427
□ correct inference	58.1	592
Total	100.0	1019

■ incorrect inference
□ correct inference

German: second vowel identical
Adj. P-value=0.000, Chi-square=60.695, df=2

if one of the two cognates lacks a second C or V; no

yes

Node 1

Category	%	n
■ incorrect inference	26.9	107
□ correct inference	73.1	291
Total	39.1	398

Node 2

Category	%	n
■ incorrect inference	51.5	320
□ correct inference	48.5	301
Total	60.9	621

English: second vowel identical
Adj. P-value=0.000, Chi-square=26.465, df=2

English: first vowel identical
Adj. P-value=0.000, Chi-square=29.983, df=1

if one of the two cognates lacks a second C or V

no; if both forms lack a second C or V

yes

no

Node 3

Category	%	n
■ incorrect inference	35.1	87
□ correct inference	64.9	161
Total	24.3	248

Node 4

Category	%	n
■ incorrect inference	13.3	20
□ correct inference	86.7	130
Total	14.7	150

Node 5

Category	%	n
■ incorrect inference	40.1	120
□ correct inference	59.9	179
Total	29.3	299

Node 6

Category	%	n
■ incorrect inference	62.1	200
□ correct inference	37.9	122
Total	31.6	322

English: first consonant or consonant cluster identical
Adj. P-value=0.000, F=17.041, df1=1, df2=320

<= identical

> identical

Node 7

Category	%	n
■ incorrect inference	50.7	72
□ correct inference	49.3	70
Total	13.9	142

Node 8

Category	%	n
■ incorrect inference	71.1	128
□ correct inference	28.9	52
Total	17.7	180

German: first consonant or consonant cluster identical
Adj. P-value=0.000, F=60.138, df1=1, df2=178

<= identical

> identical

Node 9

Category	%	n
■ incorrect inference	53.2	59
□ correct inference	46.8	52
Total	10.9	111

Node 10

Category	%	n
■ incorrect inference	100.0	69
□ correct inference	0.0	0
Total	6.8	69

Figure 9: Reading comprehension (good inferers)

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.
p. 22

rank of constraint	good inferers			bad inferers		
	consonant vowel Levenshtein	EN GE	expected or unexpected relationship	Consonant Vowel Levenshtein	EN GE	expected or unexpected relationship
1	V2	GE	unexpected	V1	EN	expected
2a	V2	EN	--- (vowel never identical)	V2	EN	expected
2b	V1	EN	expected	C2	EN	unexpected
3a	C1	EN	expected	V2	GE	unexpected
3b				V2	GE	unexpected
3c				V1	GE	expected
4a	C1	GE	expected	V2	EN	unexpected
4b				C2	GE	expected
4c				C2	GE	expected
4d				C2	GE	unexpected
5				C1	EN	expected

Table 5: Important structural characteristics in reading comprehension

Table 5 can again be used for a tentative exploration of structural constraints governing the interlingual inferencing task in the reading condition.

1. Levenshtein distance above or below the mean does not predict correct inferencing in a measurable way.

2. Of the 16 statistically significant variables that enter the model and that are listed in Table 5, 9 relate to vowels and 7 to consonants.

3. Half of the variables relate to German, half to English.

4. Of the 9 vowel variables, 4 (or 5, if 2a is included) have an unexpected relationship with the changes in correct inferences, i.e. difference coincides frequently with more correct inferences.

5. Of the 7 consonant variables, 2 have an unexpected relationship with the changes in correct inferences, i.e. difference coincides in two cases with more correct inferences.

6. The good inferers seem to be particularly efficient if the first part of the item is identical to English (V1 and C1 on levels 2 and 3 of the tree). On the other hand, if the English onset (C and V) are not absolutely identical, as in *kunne*, or if there is no English cognate at all, the German onset plays a crucial role: If it is identical, then the word can be recognized (*blive*), if not, the word is hardly recognized (*skulle*).

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.

2.7   *Discussion and support from qualitative data*

The results of these tree analyses are not absolutely clear, and some of the patterns uncovered are rather puzzling, in particular those which stand for higher comprehension proportions coinciding with phonological/graphematic differences. Nevertheless there are some preliminary conclusions that can be drawn from the analyses of the listening and the reading comprehension task:

In the listening comprehension condition, consonantal contrasts (or their absence) seem to be a more important predictor for successful inferencing than vowels. If vowels are concerned, comprehension can even be better in cases where they are different from the English or German counterpart. The same unexpected pattern in the vowel category can be found in the reading comprehension data. However, if consonants often play a measurable role, it is (with the exception of C2) in the expected direction, i.e. difference coincides with less successful inferencing.

Table 4 and Table 5 show a slightly higher importance of the first segments (onset), since differences in the second CV segment are less frequently listed and often in an unexpected relationship. This confirms Möller and Zeevaert's (in preparation: 9) observation that similarity or sameness of the onset is more important than in the rhyme in cognate recognition.

In the experimental data discussed in this contribution, we are not actually dealing with cases of automatic bi- or multilingual word recognition as investigated in other psycholinguistic paradigms (cf. Dijkstra (2003), Grosjean (2008)). Although there may well be an automatic component in cognate recognition, in this study we are rather interested in higher-level cognition applied by multilinguals to resolve the interlingual puzzles. Nevertheless, it seems useful to take into account the notions put forward in theoretical and empirical studies of the bi- or multilingual lexicon (cf. Cenoz, Jessner & Hufeisen 2003; Pavlenko 2009): Following the majority view in the field, we can assume that the multilingual lexicon is "an integrated lexicon that consists of a mix of words" (Dijkstra 2003: 17), although some of the models differ with respect to the architectural details regarding how words or features are associated with languages or organized in language-specific ways, and the configuration of activation and inhibition across or within layers vary across the different models. Most of these models, however, distinguish among different structural layers (feature, letter, word, language in BIA (Bilingual Interactive Activation), feature, phoneme, word in BIMOLA (Bilingual Interactive Model of Lexical Access)) which are incrementally activated depending on the visual/aural

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.
p. 24

input, and it is generally assumed that there does not seem to be such a thing as a "language switch". In our case, even if there were one, the participants are put in a situation where they are explicitly asked for a multilingual search across their lexicon. Based on the explorations of the data presented above, it seems that not all of the structural elements (letters, phonological features, phonemes, classes of phonemes such as consonants) participate in the same way in this search: In the auditory condition the results discussed above suggest that informants first rely on similarities of consonantal features. If these similarities are sufficiently strong or salient, differences in vowels are disregarded and chances are good that the correct cognate candidates will be selected. If consonantal patterns differ, particularly with respect to manner and place of articulation (cf. Table 4), cognate recognition becomes unlikely. Although the picture is not totally unambiguous in the case of the reading comprehension data, the tendency seems to be similar: the good inferers in particular rely on consonants. The overall difference between the cognate strings as measured by the Levenshtein algorithms only plays a significant role in the listening condition, and here only with respect to English. We can thus hypothesize that the pattern comparison that interlingual inferers apply is selective in nature and does not simply compare the forms in a holistic way.

The following passage from a think-aloud protocol in the reading comprehension condition illustrates this main focus on selected features:

(3) (think-aloud task, P: Swiss German native participant, F: Field worker; computer screen displays the form <skulle>, the task is carried out in Swiss German)

P: skulle - oh dasch aber es herzigs Wort

'skulle' – this is a cute word

ii - tönt e chli nach Tootechopf oder so

sounds like 'skull' or something

[long silence, 17s]

skulle

dasch sicher es Profilwort, wome gar nöd cha abläite

this surely is a profile word that cannot be derived

[12s]

kchäi Aanig, chani au säge i heb kchäi Aanig?

no idea, can I also say that I don't have a clue?

skulle schifaare – näi

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter.
p. 25

skulle ‚to ski’ - no

[3s]

skulle

F: was teichsch, was ächt das isch…

   what do you think what this...

P: rollen, aber nur wils zwäi -ll- hät, aber weiss au nöd

   ‚to roll’, but only because there are two –ll- but I don't know

aber da macht aso das isch absoluut…

but this is totally...

skulle

skill, ah villicht „fähig sein" oder so oder "wissen"

skill, ah, maybe 'capable of' or so, or 'to know'

skulle

ja sägemer wissen

yes, let's say ‚to know'

vo irgendwie skill, Fähigkeit oder so

from skill somehow, ability or so

The example illustrates how the participant varies the vowels (/u/, /i/, /o/) while explicitly mentioning associations between forms that share consonantal elements (e.g. *roll* because of the shared <ll> graphemes). The item is also an example that illustrates the difficulty of linking cognates that differ with respect to (onset) consonants.

However, it is by no means impossible to overcome consonantal 'impediments' (see example 1b): In retrospect, it seems that the results discussed based on Table 4 (consonant identity being more important for bad inferers than for good inferers) are in accordance with the abductive strategies that can be observed with particularly gifted multilingual subjects: The search for cognates does not stop after a first – maybe too simple and too direct – string analogy has been identified: (blive→to believe). Good inferers thus consider not only modifications of vowels, but they also search for potential rules in the consonantal domain, as the Spanish grapheme-phoneme rule discussed in example 1b. As Möller (in preparation) has shown, there are consistent patterns of interlingual plausibility that multilinguals display when they are asked to provide potential cognate forms within the Germanic languages.

3. Conclusions

Based on the analyses in Section 2, we can identify at least two partial cognitive processes that contribute to good guessing capacities based on the intra- and interlingual competences:

1) A flexible and selective comparison of features and patterns, focusing on consonants and neglecting or systematically varying the vowels.

2) Good hunches regarding when to continue and when to stop searching.

The first process can be seen as a form of perceptual tolerance in pattern recognition, i.e. the patterns are aligned in a flexible way allowing for more or less systematic variation. Coming back to the results reported in Section 2, we can now try to understand the differences between groups with different multilingual profiles: As we have seen, the most efficient interlingual inferers are those multilinguals who master two languages that are closely related to the unknown target language. This finding as such is not overwhelmingly surprising, since one could argue that these multilinguals simply have more potential transfer bases that nurture the inferencing task. In the light of the results discussed in this Section, however, I hypothesize that there is more to this than the simple number of potentially transferrable forms. As schematically represented in Figure 10, inspired by the revised hierarchical model (Kroll & Stewart 1994), I would like to suggest that multilinguals with such highly developed dia-systems – be they due to proficiency in two Romance languages or to proficiency in a Germanic dialect and a standard language – develop a meta-system that is a form of abstraction over the bilingual mental corpus. A cognate in an unknown language will first be associated with one or more than one word forms in the multilingual lexicon, and – at least if our hypothesis is correct – in the case of a bilingual Romance lexicon, as in Figure 10, the form corresponds to an abstracted meta-form that combines the features that are similar or the same in the two Romance languages and has placeholders where the two languages differ.

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter. p. 27
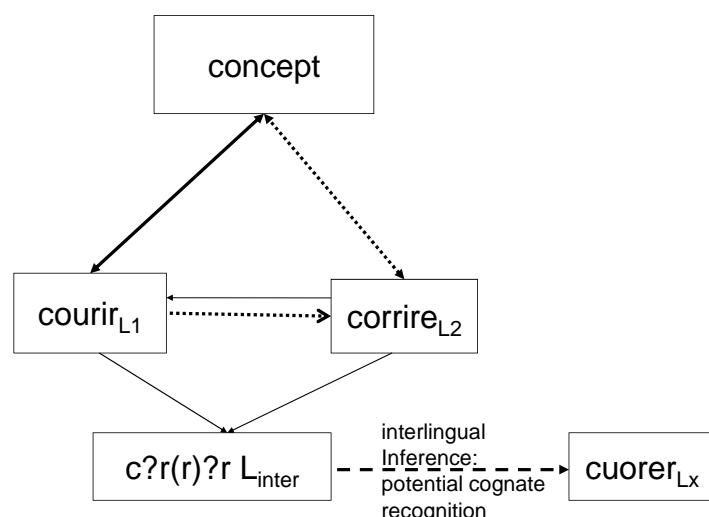
Figure 10: Schematic abstraction over French and Italian cognates of the verb 'to run' in multilinguals. <?> stands for a phoneme/grapheme placeholder. *Cuorer* is the Romontsch Sursilvan cognate.

The relative invariability of consonants and the 'unreliable' nature of vowels is a feature well known from dialectology: Schmeller's (1872) dictionary of Bavarian dialects has introduced a consonant skeleton organizing principle, since the synchronic vowel instability (caused by diachronic changes) makes an alphabetic order of a multi-dialect dictionary impossible. The finding reported in Section 2.2, i.e. that it is mainly speakers of dialect (and Standard German) who are good at interlingual inferencing, is therefore not very surprising.

This abstract schematic entry in the multilingual lexicon is the basis for the enhanced potential lexicon as measured by the interlingual inferencing task: The multilinguals develop quick and efficient interlingual heuristics within a language family regarding potential interlingual differences and similarities. These schematic forms and the heuristics that at the same time allow for their emergence and cognitive entrenchment enable multilinguals to be particularly efficient in recognizing cognates in unknown languages.

The second process listed above, at first sight, is more mysterious. How do our subjects decide when to stop and when to continue searching? Good stopping rules, according to Gigerenzer (2007), are important in simple heuristics in general just as much as in interlingual inferencing: As shown in the transcript above, sometimes searches are merely a waste of time. In other cases, however, the search is aborted too quickly since a supposed cognate has been identified that satisfies the interlingual comparison (e.g. *blive→to believe*). The data discussed here do not provide evidence that permits tapping into this particular part of the interlingual inferencing process with sufficient ecological validity. In naturalistic situations, there would

be additional, contextual information available that contributes in significant ways to this decision, most notably the semantic fit of the item within the co- and context. The artificial and decontextualized nature of the word list data investigated here deliberately neglects these important constraints, in order to focus on the purely formal, linguistic side of the process. The latter certainly plays a role in naturalistic exolingual contexts, where multilinguals are confronted with items or texts in poorly mastered or unknown languages or varieties, and we can hypothesize that the better the interlingual heuristics on the linguistic level, the better the chances to successfully understand utterances and texts in related unlearnt languages.

The challenge of the purely linguistically based process of abduction as the one described in Section 2 is therefore a dialectic one: If the subject is confronted with a cognate form, it is good to speculate based on the linguistic knowledge that is at hand. However, speculation can go too far or be a waste of time, so it is good to know what a likely candidate is, or, if none can be found that satisfies the expectations, it is important to know when to abort the search (and, in naturalistic contexts again, when it is appropriate and important to consult a dictionary or ask a proficient speaker). As the quite fuzzy description of these two processes makes clear, we are dealing with speculative and probabilistic processes that are not amenable to direct scientific observation.

The main result of the empirical studies presented in this contribution is that the quality of these interlingual guessing procedures depends on many factors, both concerning the multilingual subjects and the linguistic contrasts involved. Quality increases with increasingly proficient multilingual systems, and I have argued that the main process that enables multilinguals to be efficient in the tasks is a form of multilingual abduction, an inferencing process that exploits mutlilinguals' knowledge of what are likely and of what are unlikely correspondences across and within languages. And finally, if the tendencies discussed in Section 2.5 and 2.6 turn out to be validated by other studies, the sound correspondence section of the EuroCom materials could and should be simplified and shortened by at least 50% to the most important consonant correspondences.

4. References

Beijering, Karin, Charlotte Gooskens & Wilbert Heeringa. 2008. Predicting intelligibility and perceived linguistic distance by means of the Levenshtein algorithm. In Marjo van Koppen & Bert Botma (eds.), *Linguistics in the Netherlands*, 13-24. Amsterdam: John Benjamins.

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter. p. 29

Berthele, Raphael. 2008. Dialekt-Standard Situationen als embryonale Mehrsprachigkeit. Erkenntnisse zum interlingualen Potenzial des Provinzlerdaseins. In Klaus J. Mattheier & Alexandra Lenz (eds.), *Dialektsoziologie / Dialect Sociology / Sociologie du Dialecte. Sociolinguistica, Volume 22*, 87-107. Tübingen: Niemeyer.

Berthele, Raphael & Amelia Lambelet. 2009. Approche empirique de l'intercompréhension : répertoires, processus et résultats. *LIDIL*, 151-62.

Campbell, Lyle. 2004. *Historical linguistics : an introduction*, 2nd edn. Edinburgh: Edinburgh University Press.

Carton, Aaron S. 1971. Inferencing: a process in using and learning language. In Paul Pimsleur & Terence Quinn (eds.), *The psychology of second language learning. Papers from the second international congress of applied linguistics, Cambridge, 8-12 September 1969*, 45-58. Cambridge: University Press.

Cenoz, Jasone. 2003. The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism 7/1*, 71-87.

Cenoz, Jasone, Ulrike Jessner & Britta Hufeisen. 2003. *The Multilingual Lexicon*. Dordrecht: Kluwer Academic Publishers.

Chomsky, Noam. 1968. *Language and mind*. New York ; Chicago [etc.]: Harcourt Brace & World.

Cummins, Jim. 2000. *Language, power and pedagogy : bilingual children in the crossfire*.

de Angelis, Gessica. 2007. *Third or additional language acquisition*. Clevedon/Buffalo/Toronto: Multilingual Matters.

Dijkstra, Ton. 2003. Lexical processing in bilinguals and multilinguals. In Jasone Cenoz, Ulrike Jessner & Britta Hufeisen (eds.), *The Multilingual Lexicon*, 11-26. Dordrecht: Kluwer Academic Publishers.

Eco, Umberto. 1984. *Semiotica e filosofia del linguaggio*. Torino: Einaudi.

Gass, Susan & Larry Selinker. 1994. *Language Transfer in Language Learning*. Amsterdam: John Benjamins.

Gigerenzer, Gerd. 2007. *Gut feelings : the intelligence of the unconscious*. New York: Viking.

Grosjean, François. 2008. *Studying Bilinguals*. Oxford: Oxford University Press.

Heeringa, Wilbert, Peter Kleiweg, Charlotte Gooskens & John Nerbonne. 2006. Evaluation of String Distance Algorithms for Dialectology. *Proceedings of the Workshop on Linguistic Distances. Sidney, Australia Association for Computational Linguistics*, 51-62.

Herdina, Philip & Ulrike Jessner. 2002. *A Dynamic Model of Multilingualism. Perspectives of Change in Psycholinguistics*. Clevedon et al.: Multilingual Matters.

Hufeisen, Britta & Nicole Marx. 2007. *EuroComGerm - Die sieben Siebe. Germanische Sprachen lesen lernen*. Aachen: Shaker Verlag.

Jessner, Ulrike. 1999. Metalinguistic Awareness in Multilinguals. Cognitive Aspects of Third Language Learning. *Language Awareness Vol. 8: 3&4*, 201-9.

Kroll, J. F. & E. Stewart. 1994. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language 33*, 149-74.

Lado, Robert. 1957. *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan.

Le Pichon Vorstman, Emmanuelle, Henriette De Swart, Viktorija Ceginskas & Huub Van Den Bergh. 2009. Language learning experience in school context and metacognitive awareness of multilingual children. *International Journal of Multilingualism 6*, 258 - 80.

Levenshtein, Vladimir I. 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady, 10(8)*, 707-10.

Berthele, R. (2011). On abduction in receptive multilingualism. Evidence from cognate guessing tasks. Applied Linguistics Review, pp. 191-220. Berlin, New York: de Gruyter. p. 30

Mc Mahon, April M. S. 1994. *Understanding language change*. Cambridge:

McCann, William J., Horst G. Klein & Tilbert D. Stegmann. 2003. *EuroComRom - How to read all the Romance languages right away. 2nd revised edition.* Aachen: Shaker.

Meißner, Franz-Joseph. 2001. Vom induktiven zum konstruktiven Lehr- und Lernparadigma. Methodische Folgerungen aus der mehrsprachigkeitsdidaktischen Forschung. In Franz-Joseph Meissner & Manfred Reinfried (eds.), *Bausteine für einen neukommunikativen Französischunterricht*, 21-50. Tübingen: Narr.

Möller, Robert. in preparation. Ähnlichkeits-Intuitionen bei der Erkennung von germanischen Kognaten.

Möller, Robert & Ludger Zeevaert. in preparation. "Da denke ich spontan an Tafel" – Zur Worterkennung in verwandten germanischen Sprachen. *ZIFF*,

Müller-Lancé, Johannes. 2003. *Der Wortschatz romanischer Sprachen im Tertiärsprachenerwerb. Lernerstrategien am Beispiel des Spanischen, Italienischen und Katalanischen*. Tübingen: Stauffenburg.

Odlin, Terence. 1989. *Language Transfer: cross-linguistic influence in language learning.* Cambridge: CUP.

Pavlenko, Aneta. 2009. *The bilingual mental lexicon : interdisciplinary approaches* (Bilingual education & bilingualism [70]). Buffalo (N.Y.): Multilingual Matters.

Peirce, Charles Sanders. 1931. *The Collected Papers Vol. V.: Pragmatism and Pragmaticism.* Cambridge, MA: Harvard University Press.

Schmeller, Johann Andreas. 1872. *Bayerisches wörterbuch von J. Andreas Schmeller*, 2. edn. München,: R. Oldenbourg.

Studer, Thomas. 2008. *Inferenzen als Prinzip des Sprachverstehens* (218 Bl.). Zürich: Universität Zürich.

Wolff, Dieter. 1994. Der Konstruktivismus: Ein neues Paradigma in der Fremdsprachendidaktik? *Die neuen Sprachen, 93/5*, 407-29.