Assessing Lexical Alignment in Spontaneous Direction Dialogue Data by Means of a Lexicon Network Model

Alexander Mehler[1], Andy Lücking[2], and Peter Menke

[1] Goethe University Frankfurt am Main
[2] Bielefeld University
mehler@em.uni-frankfurt.de; {andy.luecking,peter.menke}@uni-bielefeld.de

Abstract. We apply a network model of lexical alignment, called *Two-Level Time-Aligned Network Series*, to natural route direction dialogue data. The model accounts for the structural similarity of interlocutors' dialogue lexica. As classification criterion the directions are divided into effective and ineffective ones. We found that effective direction dialogues can be separated from ineffective ones with a hit ratio of 96% with regard to the structure of the corresponding dialogue lexica. This value is achieved when taking into account just nouns. This hit ratio decreases slightly as soon as other parts of speech are also considered. Thus, this paper provides a machine learning framework for telling apart effective dialogues from insufficient ones. It also implements first steps in more fine-grained alignment studies: we found a difference in the efficiency contribution between (the interaction of) lemmata of different parts of speech.

1 Motivation

According to the *Interactive Alignment Model* [1, *IAM*], mental representations of dialogue partners on all linguistic levels become more and more similar, i.e. *aligned*, during their communicative interaction. Since the linguistic levels – phonetic, lexical, syntactic, semantic, situation model – are interconnected, alignment propagates through these levels. Via this spreading of alignment, global alignment, that is, alignment of situation models, can be a result of local alignment on lower levels. Thus, the IAM provides an account to the ease and efficiency of dialogical communication beyond explicit negotiation. Part of the efficiency of communication is the fulfillment of the dialogue task or purpose. Consequently, we would expect that *more aligned dialogues are more successful* – a proposition we make productive below.

The central mechanism that is acknowledged within the IAM is *priming*.[3] Priming is typically understood and modeled as spreading activation in neural networks. Two varieties of activation have to be distinguished:

[3] But see [2] for an argument that priming cannot be the process that implements alignment.

1. A linguistic form /x/ activates its corresponding mental representation x within the interlocutors. We simply call this *activation*.
2. A representation y, which is activated by a form /y/, also activates representations which are related to y. The kind of relation depends on the linguistic level of which y is an element. For example, if y is the phonological representation of *can*, the phonologically similar representation *pan* may be co-activated. Since the contents of many cans can be heated in pans, the semantic representations of both forms also trigger each other. The mediated activation of a representation x by a form /y/ is termed *co-activation*.

The linguistic forms produced and perceived in a dyad do not only prime their corresponding representations, they also co-activate a set of related representations. A model that captures the structure of dialogue lexica of speakers is a network model of interlinked nodes. The nodes of this model represent linguistic elements of a certain kind. Since we are concerned with lexical alignment in this paper, the nodes in our model represent lemmata. In order to give an impression of the phenomena we are interested in, consider the following score of a dialogue extract:[4]

A:	*street lights*		*lamps*
B:		*street lamps*	

A and *B* talk about the same (plural) referent, what we will call the *topic* of a contribution. The term *A* proposes (*street lights*) is corrected by *B* (*street lamps*). *B*'s correction is then partly taken up by *A* (*lamps*). From the perspective of alignment, the dialogue lexica of the interlocutors contain three related nouns which are linked among each other in corresponding ways. The interlocutors finally align on the repeated use of a certain noun, namely *lamp*.

Observable evidence for alignment like the *lamp* example is ubiquitous in human communication. This notwithstanding, a correlation between the type of communication and extent of alignment has been reported. [4] found that speakers in a task-oriented dialogue setting are more receptive to priming than speakers in a spontaneous dialogue setting. The authors used common linear regression as the statistical analysis tool. Recently, [5] developed a network-based framework to model alignment in dialogue, the so-called *Two-Level Time-Aligned Network* (TiTAN) model. The TiTAN model has already been applied to strictly task-oriented dialogue data [6]. In this paper, we use the TiTAN model to assess alignment in more spontaneous dialogue data, namely direction dialogues. We do that by classifying dialogues for being effective or ineffective according to their main function, that is, direction giving.

In the following Section 2 we shortly point out two shortcomings of previous approaches to measure priming or alignment which are overcome by the TiTAN model introduced in Section 3. After that, the TiTAN model is applied to natural

[4] The extract is taken from dialogue no. 24 around second 600 of the collection from [3] – see Section 4 for some more details. In its German original form, the sequence of nouns is *Straßenlampen* – *Straßenlaternen* – *Laternen*.

language direction data. The data and the results are described in Section 4, which is followed by a conclusion that summarizes our findings.

2 Related Work

The approach followed here diverges in two respects from related work that tries to measure priming or alignment.

The earliest work on assessing alignment-related properties of (written) texts in quantitative terms is the lexical adaption model proposed by [7]. In a nutshell, Church measured the frequency of primed words in comparison to unprimed ones in the second half of split documents. A related measurement of the recurrence of syntactic patterns was conducted by [4], who account for the repetition of phrase structure rule instances within the Switchboard [8] and the HCRC Map Task [9] corpora.

A priming assessment that relates counting repeated elements to task achievement was implemented by [10]. They trained a *Support Vector Machine* (SVM) to predict task success from lexical and syntactic repetition in the HCRC Map Task corpus. Thus, the study is also precursor for the efficiency of aligned dialogue hypothesis pursued in the empirical part of this paper. The SVM is applied to time stamps in the data, indicating the proportion of variance that can be explained by the model.

The accounts for assessing priming effects in natural language data so far underlie two restrictions:

1. They focus on the *repetition* of elements, that is, they do not account for co-activation and linked representations.
2. They operate on fairly *arbitrary temporal units* that were artificially imposed on the data.

The model proposed by [5], the *Two-Level Time-Aligned Network* (*TiTAN*) model, avoids both afore-mentioned restrictions. The temporal units which carry the alignment process are dialogue turns, genuine components of conversations. The network structure allows for capturing co-activation of related elements. The next section explains how the TiTAN model of direction givings looks like.

3 Modeling Dialogue Lexica as TiTAN Series

During their conversation, interlocutors establish a so called *dialogue lexicon* [1] of commonly or differently used words. On the one hand, they may reuse words that their partner used the same way or at least similarly within their conversation. Alternatively, interlocutors may use the same words but for different things or may introduce new words that were not used before. Sameness (and conversely difference) of word usage, thus, is detected according to the extensional criterion of aboutness. What words are about is called their *topic* in the following. By speaking about similar word usages we refer to the similarity of the lexical contexts of words [11]. In the present scenario, this context is identified with the basic structural unit of dialogues, that is, the turn [12].

From this point of view, the generation of a dialogue lexicon is conceived as a process in which a lexical network grows turn by turn based on the word usages of the dialogue partners. In such a *dialogue lexicon network*, vertices denote lexical items while the strength of their edges denote the number of contexts in which these words co-occurred until the corresponding point in time. In this sense, the generation of a dialogue lexicon appears as a *time series* that emits lexical networks at its different time points. It is this combined notion of complex networks [13] and time series that is used to model the build-up of dialogue lexica as a result of dialogical communication. In this section, we briefly recapitulate this model in terms of so called *Two-Layer Time-Aligned Network* (TiTAN) series [5] and introduce its instantiation in the context of direction givings.

Generally speaking, a TiTAN series is a time series $\{L_t \,|\, t \in \mathbb{N}\}$ of indexed graphs L_t that model the dialogue lexicon of a dyadic conversation at time t. Each of these graphs L_t is partitioned into two layers, A and B, representing each interlocutors' part of the dialogue lexicon. In order to instantiate the notion of a TiTAN series in the framework of direction givings, we start with formalizing dialogue lexica before we explain how TiTAN series are serialized.

Formally speaking, the dialogue lexicon of a dyadic conversation among two interlocutors A and B at time t is modeled as a labeled graph $L_t = (V, E_t, \mathcal{L})$. In this graph, the vertex set V is partitioned into non-empty disjunct subsets V_A and V_B whose elements denote the words used by interlocutor A and B, respectively, to perform the task of direction giving. The vertices in V are labeled by the surjective function $\mathcal{L} : V \rightarrow L_V$ where, in our case, the set of labels L_V consists of lemmata. Analogously, the edge set E_t is partitioned into three non-empty disjunct subsets E_{AB_t}, E_{A_t}, E_{B_t} where all edges $\{v, w\} \in E_{AB}$ end at vertices $v \in A, w \in B$, while all edges $\{x, y\} \in E_X$, $X \in \{A, B\}$, end at vertices $x, y \in V_X$. E_A and E_B capture intrapersonal lexical relations, while edges in E_{AB} are used to link lexical items shared among the interlocutors. The subgraphs $L_{A_t} = (V_A, E_{A_t}, \mathcal{L})$ and $L_{B_t} = (V_B, E_{B_t}, \mathcal{L})$ are called the A- and B-layer, respectively, of the two-layer graph $L_t = (V, E_t)$ at time t. They are denoted by the projection functions $\pi_A(L_t) = L_{A_t}$ and $\pi_B(L_t) = L_{B_t}$. In terms of our application area, layer A represents the dialogue lexicon of interlocutor A, layer B represents the dialogue lexicon of interlocutor B, while the graph L_t provides a unified model of their overall dyadic dialogue lexicon.

The networks defined so far model linguistic units and their relations. However, they do not distinguish between seldomly and frequently intantiated relations. This asymmetry is accounted for by assigning weights to the edges in L_t. Thus, dialogue lexica are modeled as weighted labeled graphs $L_t = (V, E_t, \mu_t, \mathcal{L})$ that are indexed by the point in time $t \in \mathbb{N}$ at which they are spanned. Recall that t is derived from the dialogue turns of the interlocutors and, thus, from a dialog-inherent time-related ordering. In this sense, a TiTAN series is serialized according to the contributions of the interlocutors manifested and organized as turns. As a two-layer graph, L_t is divided into the subgraphs

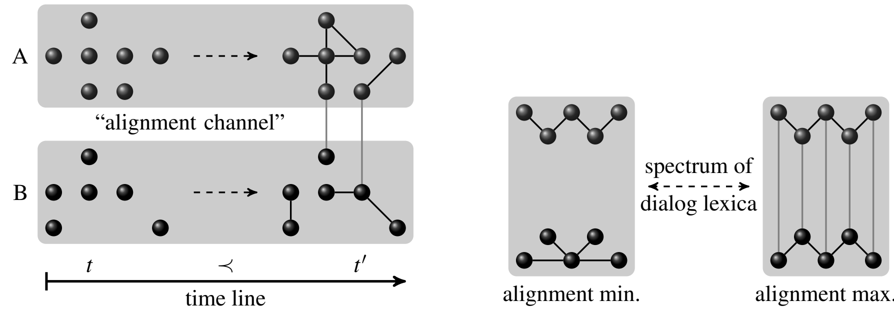
$\pi_A(L_t) = L_{A_t} = (V_A, E_{A_t}, \mu_{A_t}, \mathcal{L})$ and $\pi_B(L_t) = L_{B_t} = (V_B, E_{B_t}, \mu_{B_t}, \mathcal{L})$ according to the distribution of L_t over the agents A and B at time t. The spanning of edges within L_t is done as follows [5, p. 1453]:

- *Intrapersonal links:* if at time t, agent $X \in \{A, B\}$ uses a word form as an instance of lemma $l \in L_V$ to express the current turn's topic $T = T(t)$, intrapersonal links between vertex $v \in V_X, \mathbb{L}_X(v) = l$ are generated, and all other vertices w in layer L_{X_t} whose lemma $\mathcal{L}(w)$ was used by X in the same or any preceding turn to speak about the same topic T. If any of these edges $e = \{v, w\}$ already exists, its weight is incremented by 1, that is, $\mu_t(e) = \mu_{t-1}(e) + 1$. Initially, all edges have a weight of 1.
- *Interpersonal links:* if at time t, agent $X \in \{A, B\}$ uses a word form as an instance of lemma $l \in L_V$ to express the topic $T = T(t)$, which has been expressed by the dialogue partner $Y \neq X$ in any preceding turn on the same topic by means of the same lemma, an interpersonal link $\{v, w\} \in E_t$ between $v \in V_A$ and $w \in V_B$ is generated for which $\mathcal{L}(v) = \mathcal{L}(w) = l$, given that this link does not already exist. Otherwise, its weight is increased by 1.

With the passing of time, this process generates a series of dialogue lexica L_t that are indexed by the corresponding time point t. Figure 1(a) provides a schematic visualization of this construction process of a TiTAN series. The starting point is given by completely unlinked dialogue lexica L_{A_0} and L_{B_0} of the interlocutors A and B. Following the afore-given construction procedure, the lexica are networked turn-wise by adding intra- and interpersonal links. A TiTAN, thus, allows for modeling for each turn the degree of structural coupling of the dialogue lexica of both interlocutors. It finally results in a dialogue lexicon that manifests the degree of lexical alignment at the end of the conversation of both interlocutors. To see this, look at Figure 1(b), which shows two extreme values of dialogue lexica: the lower bound is given by two layers L_{A_t} and L_{B_t} of the overall dialogue lexicon L_t that are completely disconnected and internally structured in completely different ways. Such a situation occurs if both agents always use different words or denote the same topics always differently. The upper bound is set by two isomorphic layer graphs that are fully linked. This scenario corresponds to a dialogue in which both agents always use the same words the same way. Due to thematic progression of natural dialogues, constraints by stylistics and verbal economy, and psychological factors of various kinds, neither of these extremal points is to be expected to be realized by dialogical conversation. They delineate, however, theoretical boundary values that make lexical alignment a measurable property [5].

In the framework of task-related conversations like direction giving, alignment is supposed to be bound up with communicative success, i.e. efficiency [1, p. 172].

The question arises how to measure whether a dialogical interchange is efficient or not. Using TiTAN series to represent such dialogues, we hypothesize that the class of effective directions can be separated from the class of ineffective directions in terms of the topology of the final state of the dialogue lexica L_t.



(a) Illustration of a *Two-layer Time-Aligned Network* series. Initially, the lexica of interlocutors A and B are unlinked (left of dashed arrows). Dialogue turn by dialogue turn, the interlocutors' lexica are networked such that a dialogue lexicon emerges that is spanned by intra- and interpersonal links across the alignment channel (right of dashed arrows).

(b) The extrema of the graph-theoretical modeling of lexical alignment: completely independent (left) vs. fully linkage of identical lexica (right). Note that alignment minimum and maximum are theoretical extrema that are not to be expected to be found in real dialogues, which are supposed to populate the spectrum of dialog lexica.

Fig. 1: (Co-)Activation of representations within the dialogue networks of interlocutors A and B: A TiTAN series illustration (a) and structural extrema (b).

In other words, we hypothesize that the way lexical items are connected and clustered in a dialogue lexicon informs about the status of the corresponding direction giving. If this is true, it should be possible to utilize complex network theory [13] to represent dialogue lexica by topological indices that are finally input to unsupervised learning of the class of effective and ineffective directions. This is the way, we proceed in this paper. More specifically, we apply *Quantitative Network Analysis* (QNA) [14, 15] to represent and classify dialogue lexica by means of complex network theory. In the present area of application, QNA involves three steps of modeling:

1. *Quantitative graph modeling*: initially, each dialogue lexicon is represented by a vector of topological indices that model its network structure.
2. *Feature selection:* in the next step, a genetic search is performed to find salient features within the vectors that best separate effective and ineffective dialogues. Note that this process of feature selection may stop at a local maximum as it does not necessarily find an optimal feature subset.
3. *Classification:* based on the appropriately projected feature vectors, a hierarchical agglomerative clustering is performed together with a subsequent partitioning that is informed about the number of target classes. We use *complete linkage* together with the *Mahalanobis distance* to perform this step. Note that we use MATLAB to make any of these computations. Note

also that the Mahalanobis distance is used to handle correlations between features.

To sum up, QNA takes the set of input dialogues together with the parameter space of linkage methods and distance measures to find out the feature subset that best separates the data according to the underlying classification. In the present study, we utilize a subset of indices of complex network theory together with a subset of indices that were invented to model dialogue lexica [5]. See [5] and [15] for a summary of this quantitative graph model. All in all, 50 topological indices were computed per input dialogue to model its structural characteristics. Note that we exclude simple frequency oriented indices (e.g., the number of vertices or edges). In Section 4, we discuss eight instantiations of this model by experimenting with a set of 25 dialogues about directions.

4 Experimentation

4.1 Data

The speech data the TiTAN model is applied to are taken from the *Bielefeld Speech and Gesture Alignment Corpus* (SaGA) [3]. The primary data of the SaGA corpus are made up of 25 direction dialogues. After finishing a simulated bus ride through a virtual town, one participant explains the route taken and some sights passed to a second participant.

Video and audio recordings were made of the experiments, and on their basis, an orthographic transcription of speech on the level of words has been created. Typical phenomena of spontaneous speech (for example, clitics, elisions, assimilations, and spontaneous neologisms) were transcribed according to guidelines in order to ensure consistency.

These transcriptions were tagged with part-of-speech and lemma information by a system consisting of the eTagger of the eHumanities Desktop [16], a central trigram HMM tagger that has been trained on the German Negra Corpus.[5] The *Stuttgart-Tübingen Tag Set* (STTS, cf. [17]) was used, along with pre- and post-processing mechanisms that are specialized in the handling of the phenomena of spoken language mentioned above. Preprocessing methods map recurring word form variants to their standardized counterparts before tagging. Postprocessing mechanism apply several heuristics to unrecognized words that help to identify neologisms – for example those that had been constructed from two or more known words (e.g., "Peitschenlampe‘" – *whip lamp*, constructed from the nouns "Peitsche" / *whip* and "Lampe" / *lamp*). Still, there were word forms that could not be detected or handled automatically. These tokens have been manually corrected after applying the tagger.

Since we are not concerned with well-formedness or related grammatical notions, but rather with regularites of word use, the syntactically fine-grained POS of the STTS are too detailed. Thus, we mapped the STTS onto the functionally

[5] `http://www.coli.uni-saarland.de/projects/sfb378/negra-corpus/`

basic types N(oun), V(erb), ADJ(ective), ADV(erb) JUNC(tors), PREP(ositions), DET(erminers), PRO(nouns), and PART(icles). A fourth type, called REST, collects the remainder of POS like interjections and fragments. These basic categories are used in the construction of dialogue lexica.

For the construction of a TiTAN series, information about turns and their respective topic are required. As a consequence, the turn boundaries needed for the construction of a TiTAN were annotated manually for all of the 25 dialogues. The topics we acknowledge are derived from the stages of the route through the virtual town of the primary data. The SaGA town and its 12 stages (topics) are shown in Figure 2. In addition, there is a 13th topic called *SaGA* which indicates turns that are about (large parts of) the whole virtual town. The label *META* is used to classify turns that do not relate to the route, but rather negotiate discourse issues or interpersonal concerns.

Each dialogue has been rated with respect to whether the interlocutors converge on a suitable description of the SaGA route. The criterion is whether the addressee has been put into the position to find the way from the sculpture to the fountain without going astray. We distinguish three cases or classes: 1. The direction is correct; 2. The direction is partially correct, but sufficient for the purpose to cross the SaGA town; 3. The direction is full of holes and useless. If, for instance, a participant mistakes the conifers in the park for leaf trees but apart from that gives a right direction, the dialogue is classified into the second class. Class 1 and 2 are grouped together into "correct" directions. In sum, there are 17 wrong and 8 correct directions. For each of these 25 dyads, a separate dialogue lexicon network has been built according to Section 3.

One might object that the occurrence of alignment is independent from the validity of the given direction. Note, however, that the classification of dialogues is not concerned with their correctness in the first place. In particular class 2 above accounts for directions that are false strictly speaking, but nonetheless carry enough information to let the addressee find the way. Finally, the class 3 dialogues are clearly faulty. *So what is the root of the matter?* [1, p. 172] emphasize that "alignment of situation models is central to successful dialogue". No matter whether the situation models are correct or not, a precondition is that the dialogue participants have situation models at all! There are reasons to assume that this is the problem with class 3 dialogues. The direction givers' models of the SaGA town are fragmentary – the models contain gaps. It is questionable whether fragmentary models can be conceived of as situation *models* at all. As it stands, we are aware of these theoretical obstacles, but regard our classification approach as feasible.

4.2 Evaluation

In this section, we describe the experimental scenarios by which we test our classification hypothesis introduced in Section 3. This hypothesis says that the efficiency of a direction giving in dialogical communication can be detected based on the topology of the final state of the corresponding dialogue lexicon. As described in Section 3, we test this hypothesis in the framework of *Quantitative*

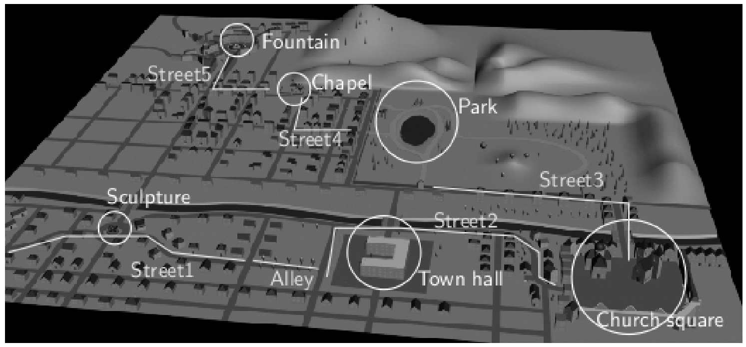


Fig. 2: Overview of the virtual SaGA town, with topics marked.

Network Analysis (QNA). More specifically, we test 8 different variants of this hypothesis (as summarized in Table 1):

- *Variant* [N]*:* we start with considering nouns only. The idea behind this approach is that nouns are mainly used by interlocutors to refer to the reference universe of the direction giving – see the entities of the virtual SaGA town marked in Figure 2. We expect that the efficiency of directions is more easily identified by means of the nominal subnetworks of the corresponding dialogue networks.
- *Variant* [$N|A$]*,* [$N|V$] *and* [$N|V|A$]*:* alternatively, we experiment with additionally considering adjectives and verbs. The reason to take these POS into account is that many of their instances have a descriptive meaning in relation to the reference universe of the direction giving (as, e.g., the verb *to turn* in *Turn to the left*).

In addition to these four variants, we consider those subnetworks that exclude words with a meta-communicative function (see the Rows 1–4 and the column *Meta*, which codes whether words are included that are tagged by *META* according to Section 4.1). These are words (as, e.g., *to think* in *Let me think*) that do not have a referential meaning regarding the reference universe of the direction giving, but serve, for example, to organize the dialogue. In our corpus of 25 direction givings, we have annotated 5,561 word forms with a meta-communicative function in relation to the overall set of 45,190 word forms that we were manually annotated. Thus, more than 10% of the word forms were used for meta-communicative reasons. From this perspective, one may expect an effect of excluding or including this class of words.

Table 1 summarizes the results of our findings. It shows that the best performing variant is based on selecting nouns without any meta-communicative function (see Row 1). This variant produces an F-score of more than 96%. The F-score (or F-measure) is the harmonic mean of recall and precision of the computed classification in relation to the correct classification of the data into 17 ineffective and 8 effective dialogues. An F-score of 96% means that nearly all

Table 1: Summary of the results of differently parameterized *Quantitative Network Analyses* (QNA). Row no. 9 shows the average F-score of these variants. The last column denotes the number of features output by the genetic search of the best performing subset of features as part of QNA (see Section 3).

No.	**Setting**	**Meta**	**Procedure**	**F-score**	**Features**
1.	$[N]$	yes	QNA	.96057	16 / 50
2.	$[N\|V]$	yes	QNA	.92	18 / 50
3.	$[N\|A]$	yes	QNA	.91651	20 / 50
4.	$[N\|V\|A]$	yes	QNA	.88171	21 / 50
5.	$[N]$	no	QNA	.92194	21 / 50
6.	$[N\|V]$	no	QNA	.87771	18 / 50
7.	$[N\|A]$	no	QNA	.88171	22 / 50
8.	$[N\|V\|A]$	no	QNA	.91651	24 / 50
9.	average over non-random approaches			.9096	20
10.	random baseline known-partition			.58668	
11.	random baseline equi-partition			.58583	

dialogues have been classified correctly. If we additionally consider verbs, the F-score decreases to 92% (Row 2). The loss of classification is even higher if we separately consider the network of adjectives and that of verbs and adjectives (Row 3 and 4). Thus, although adjectives and verbs have denotational meanings in the dialogues analyzed here, they do not help to separate the class of effective and ineffective direction givings to the same degree as nouns only.

These results seem to be contra-intuitive. Denotations of orientations and movements should be key ingredients of a successful direction giving. However, they are relational in character as they depend on the things they relate. Regarding situation models, a precondition for relational specification is that the objects in question are (correctly) spread out on the mental model. This in turn requires that the objects are available to the interlocutor. Objects are typically denoted by nouns or noun phrases. If the direction giver can name the things he wants to talk about, he can relate them to each other or to the direction follower. Thus, correct $[N|V|A]$-dialogues depend on correct $[N]$-dialogues. Besides this logical relationship, however, verbs, and adjectives may be the source for errors beyond nominal expressions. The decreasing F-score of $[N|V|A]$, $[N|V]$, and $[N|A]$ variants in comparison to the $[N]$ variant is very probably due to the asymmetrical status of the $[N]$ partitions of the dialogues in relation to their adjective- and verb-based partitions.

What happens if we additionally consider words with meta-communicative functions? As shown by the rows 5 through 8 in Table 1, there is a negative effect of including meta-communicative words. However, the differences being observed are rather marginal so that we conclude that there is only a small effect of either including or excluding this class of words. Meta-communicative acts typically provide information that the addressee has either understood the direction or that he could not follow. Thus, meta-communicative turns are used to convey a sort of binary information. As this information does not relate to the direction proper, it may be the reason for the lack of classificatory power being observed.

In order to further assess the quality of our results, we computed two random baselines (Row 10 and 11 in Table 1): the baseline called *known-partition* has information about the number of instances of the target categories. That is, by knowing that there are 17 ineffective and 8 effective dialogues, this baseline randomly generates two subsets of these cardinalities to compute the corresponding F-score. By repeating this procedure 1,000 times, we get an expected F-score of about 58%. This score is a little bit smaller if we consider the second random baseline that assumes equal sizes of the target categories (in our case 12 and 13). Obviously, all topology-related classifiers clearly outperform these two baselines. Thus, we can conclude, at least until any future falsification, that the efficiency of a direction giving is encoded into the *structure* of its dialogue lexicon.

5 Conclusion

In this paper, we applied *Two-Layer Time-Aligned Network* (TiTAN) series in the context of direction givings. Based on this graph model, we implemented several classifiers that solely explore the structure of dialogue lexica to assess their efficiency. By example of a corpus of 25 dialogues, we have shown that topological indices of dialogue lexica can indeed reveal this status. We also observed that lexical units with meta-communicative functions have a small effect on classification. This is in support of the observation that lexical manifestations of dialogue organization have a some impact on the efficiency of direction givings. Furthermore, we observed that the networking of nouns has the highest classificatory power, while the subnetworks of adjectives and verbs are less informative. One reason for this finding may be the outstanding referential meaning of nouns in conjunction with their semantic specificity. There are several POS that we did not consider here. Apart from adverbs, this relates to instances of closed POS. In future work will consider these classes and their role in the organization of dialogue lexica too.

Acknowledgements

This work has been supported by the German Research Foundation (DFG), which granted the Collaborative Research Center 673, and the German Ministry of Education and Research, which funds the *Linguistic Networks* project.

References

1. Pickering, M.J., Garrod, S.: Toward a mechanistic psychology of dialogue. Behavioral and Brain Sciences **27**(2) (2004) 169–190

2. Schiller, N.O., de Ruiter, J.P.: Some notes on priming, alignment, and self-monitoring. Behavioral and Brain Sciences **27** (2004) 208–209
3. Lücking, A., Bergmann, K., Hahn, F., Kopp, S., Rieser, H.: The Bielefeld speech and gesture alignment corpus (SaGA). In: Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality, Malta, 7th International Conference for Language Resources and Evaluation (LREC 2010) (5 2010) 92–98
4. Reitter, D., Keller, F., Moore, J.D.: Computational modelling of structural priming in dialogue. In: Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics., New York, NY (2006) 121–124
5. Mehler, A., Lücking, A., Weiß, P.: A network model of interpersonal alignment in dialogue. Entropy **12**(6) (2010) 1440–1483
6. Mehler, A., Weiß, P., Menke, P., Lücking, A.: Towards a simulation model of dialogical alignment. In Smith, A.D.M., Schoustra, M., de Boer, B., Smith, K., eds.: The Evolution of Language. Proceedings of the 8th International Conference (EVOLANG8), Singapore, World Scientific (2010) 238–245
7. Church, K.W.: Empirical estimates of adaptation: The chance of two noriegas is closer to $p/2$ than p^2. In: Proceedings of Coling 2000, Saarbrücken, Germany (2000) 180–186
8. Wheatley, B., Doddington, G., Hemphill, C., Godfrey, J., Holliman, E., McDaniel, J., Fisher, D.: Robust automatic time alignment of orthographic transcriptions with unconstrained speech. In: Proc. ICASSP-92. Volume 1. (1992) 533–536
9. Anderson, A.H., Bader, M., Gurman Bard, E., Boyle, E., Doherty, G., Garrod, S.: The HCRC map task corpus. Lang. Speech **34** (1991) 351–366
10. Reitter, D., Moore, J.K.: Predicting success in dialogue. In: Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics (ACL), Praque, Czech Republic (2007) 808–815
11. Miller, G.A., Charles, W.G.: Contextual correlates of semantic similarity. Language and Cognitive Processes **6**(1) (1991) 1–28
12. Schegloff, E.A.: Sequence Oeganization in Interaction. Cambridge, Cambridge (2007)
13. Newman, M.E.J.: The structure and function of complex networks. SIAM Review **45** (2003) 167–256
14. Mehler, A.: Structural similarities of complex networks: A computational model by example of wiki graphs. Appears in: Applied Artificial Intelligence (2008)
15. Mehler, A., Pustylnikov, O., Diewald, N.: Geography of social ontologies: Testing a variant of the Sapir-Whorf Hypothesis in the context of Wikipedia. Computer Speech and Language (2010)
16. Gleim, R., Waltinger, U., Ernst, A., Mehler, A., Esch, D., Feith, T.: The ehumanities desktop - an online system for corpus management and analysis in support of computing in the humanities. In: Proceedings of the Demonstrations Session of the 12th Conference of the European Chapter of the Association for Computational Linguistics EACL 2009, Athens (3 2009)
17. Schiller, A., Teufel, S., Thielen, C.: Guidelines für das Tagging deutscher Textcorpora mit STTS. Technical report, Technical Report, IMS-CL, Universität Stuttgart (1995)