Using Hybrid Connectionist Learning for Speech/Language Analysis

Volker Weber and Stefan Wermter

University of Hamburg
Department of Computer Science
22765 Hamburg, Germany
{weber|wermter}@informatik.uni-hamburg.de

Abstract. In this paper we describe a *screening* approach for *speech/ language analysis* using *learned*, *flat* connectionist representations. For investigating this approach we built a hybrid connectionist system using a large number of connectionist and symbolic modules. Our system SCREEN[1] learns a flat syntactic and semantic analysis of incremental streams of word hypothesis sequences. In this paper we focus on techniques for improving the quality of pruned hypotheses from a speech recognizer using acoustic, syntactic, and semantic knowledge. We show that the developed architecture is able to cope with real-world spontaneously spoken language in an incremental and parallel manner.

1 Introduction

Processing real-world spontaneously spoken language in computational models causes more problems than the analysis of written texts since spontaneous language is often irregular, faulty and heterogeneous. Besides the "noise" produced by a human speaker (interjections, pauses, repetitions, repairs, restarts, etc.) there is also noise induced by the faulty output of the speech recognizer. In this article we will describe an approach which combines real-world speech and language processing based on a hybrid connectionist learning system.

Several connectionist learning techniques have been used for text analysis in the past. McClelland, Kawamoto and Miikulainen used connectionist learning for case role analysis [9, 12]. Syntactic tagging with a connectionist learning approach has been studied by Hanson and Kegl [4]. Furthermore, Elman and Mozer examined sequence processing [3, 13], McMillan, Mozer, Towell, and Shavlik dealt with rule induction [10, 15] and Wermter investigated a scanning understanding of complex phrases [25]. These connectionist approaches concentrated on *written text* while our approach in this paper deals with real-world *spontaneously spoken language* analysis. Only recently there has been more interest in speech/language integration [5, 8, 16, 17, 18]. However most approaches have not yet explored *connectionist learning* techniques of *flat* syntactic and semantic representations for improving the quality of speech/language analysis.

[1] **S**ymbolic **C**onnectionist **R**obust **E**nterpris**E** for **N**atural language

Different techniques have to be used for analyzing spoken language than for analyzing text. In particular, we consider the following general properties as important for analyzing spoken language.

Learning: Spontaneous language is often irregular. Therefore it is difficult to design rules which are able to cope with unrestricted real-world spontaneous language utterances. Therefore, we decided to use learning techniques which are able to generalize. Another advantage of a learning approach is that a system built for a specific domain can be ported easier to another domain.

Fault-tolerant processing: Connectionist learning approaches are able to cope with irregularities. Since spontaneous language contains phenomena like repairs, interjections, pauses, repetitions, false starts, ungrammatical constructions, and unforeseeable semantic constructions it is necessary to deal with them in a fault-tolerant manner. Some local irregularities can be dealt with by the inherent fault-tolerance of connectionist learning models while other often occurring irregularities can benefit from the design of specialized error-detection modules.

Flat screening analysis: Since spontaneous language utterances possess very different constructions it is difficult to design a high level interpretation which is general enough to cover all utterances. Therefore, the interpretation level should be rather close to a *flat* syntactic and semantic analysis for the integration with speech recognition since a deep interpretation level is too restrictive for faulty spontaneous language.

Based on these principles we designed and implemented the incremental and parallel hybrid connectionist architecture SCREEN for speech/language integration [23]. It uses a large number of connectionist networks but also a number of symbolic modules for rather simple tasks (like for instance the comparison of two words). Connectionist networks support learning, fault-tolerance, and flat screening analysis that we identified as being essential for the integration of speech/language analysis.

Current speech-recognizers compute word hypotheses which form a word graph. The huge number of paths through a word graph often exceeds the computational capacities for further syntactic and semantic analysis. Therefore the search space of a word graph has to be pruned. This can be done with an acoustic score computed by the speech recognizer. Unfortunately the acoustic score alone does not give enough evidence for good sentence hypotheses. Flat syntactic and semantic analysis provides additional knowledge for testing whether a given path is syntactically or semantically more plausible than another one in order to reduce the search space of good hypotheses which have to be interpreted.

SCREEN is a new hybrid connectionist system for the flat syntactic and semantic analysis of spontaneous speech. First we have explored the feasibility of dealing with spontaneous language using transcripts [20, 23]. After these preliminary successful case studies with transcripts we extended the SCREEN system to processing streams of word hypotheses generated by a speech recognizer. In this paper we describe the system architecture, give a description of the flat representations, point out the advantages of a screening analysis for improving the quality of hypotheses, and provide a detailed analysis of the running system.

2 From acoustics to flat syntactic and semantic analysis

2.1 The output of a speech recognizer

Acoustic signals are transformed to word hypotheses by a speech recognizer. Current speech recognizers produce many different word hypotheses for an acoustic signal. Typically the output of a speech recognizer is a list of word hypotheses (table 1) which represent a word graph (figure 1). The list consists of a start-time and an end-time in seconds, the recognized German word (in this paper with its literal English translation), and an acoustic score which gives the acoustic plausibility in the range from 0.0 (unplausible) to 1.0 (plausible).

The word graph can be constructed from the word hypotheses by connecting a word hypothesis with end-time X with an immediately following word hypothesis with start-time X+0.01s. For instance, the first word hypothesis in table 1 "#PAUSE#" (end-time 0.14s) can be connected with the word hypothesis "schönen" or "schön" ("fine"; start-times 0.15s). In table 1 and figure 1 we have chosen a rather small list of word hypotheses to be able to show the full word graph and to illustrate the principle of improving the quality of sentence hypotheses with flat syntactic and semantic knowledge. Usually word graphs are much bigger and the necessity for pruning is much more obvious.

start time	end time	word hypothesis		score by speech-recognizer
0.00	0.14	#PAUSE#		4.074477e-16
0.15	0.44	schönen	(fine)	8.200215e-15
0.15	0.44	schön	(fine)	8.200215e-15
0.45	0.62	bis	(until)	1.479265e-13
0.63	0.83	dann	(then)	7.858106e-17
0.63	0.83	sollen	(should)	1.900592e-17
0.63	0.84	dann	(then)	6.550429e-17
0.63	0.87	sollen	(should)	4.126975e-17
0.63	0.87	dann	(then)	8.347788e-17
0.84	1.02	drauf	(on)	1.626915e-17
0.85	0.87	<NIB>	not recognized	5.582233e-30
0.88	1.02	auf	(good)	4.482144e-17
1.03	1.57	Wiedersehen	(bye)	1.168121e-14
1.58	3.04	#PAUSE#		7.456669e-09

Table 1. Word hypotheses for sentence "Fine, until then, good bye". Note that for illustration and didactic reasons we have chosen an extremely simplified word graph with a small number of word hypotheses.

Following the edges of the word graph we can build sentence hypotheses. There are plausible hypotheses like "Schön bis dann auf Wiedersehen" ("Fine until then good bye") or syntactically and semantically implausible ones like

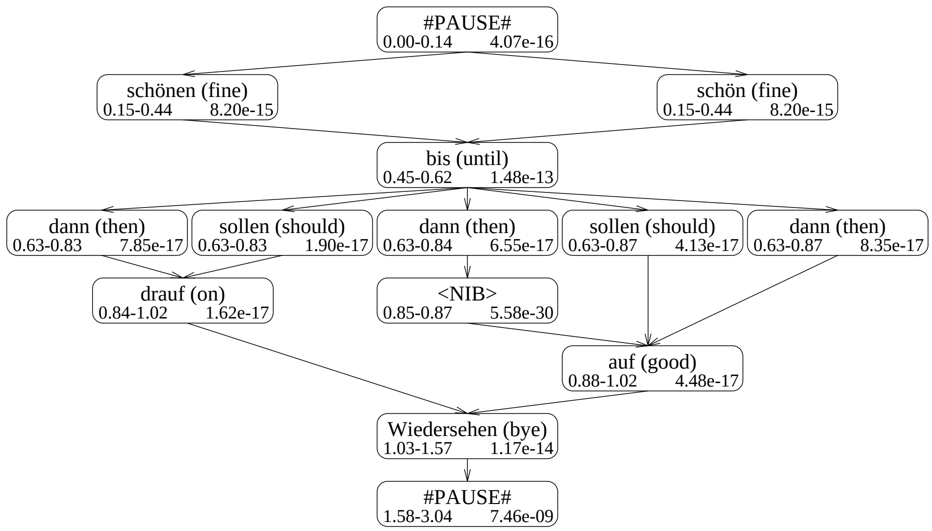

Fig. 1. Word graph for the word hypotheses from table 1.

"Schönen bis sollen drauf Wiedersehen" ("Fine until should on bye"). However both sentence hypotheses have an average acoustic score of about 2.4e-14 per word hypothesis so that the acoustic score does not give enough evidence for favoring a particular sentence hypothesis. Furthermore, this example illustrates that the syntactic and semantic processing has to be fault-tolerant to cope with such unforeseeable irregular input.

2.2 Flat syntactic and semantic analysis

For dealing with faulty spontaneous input we decided to use a flat syntactic and semantic representation (see figure 2 for a simple example of flat representations). We use 13 different syntactic categories for representing words (**N**oun, prono**U**n, ad**J**ective, **V**erb, **A**dverb, p**R**eposition, **C**onjunction, **D**eterminer, nu**M**eral, **I**nterjection, **P**ast participle, pause (/), and **O**ther) and 20 different semantic categories (statements of being (**IS**), having (**HAVE**), meeting (**MEET**), moving (**MOVE**), and **AUX**iliary; **SUG**gestions, **SEL**ections, and **UTTER**ances; **PHYS**ical, **ABS**tract, and **ANIM**ate objects; time or location points (**HERE**), **S**ou**RC**es, and **DEST**inations; **LOC**ation and **TIME**; **N**egati**O**ns, agreements (**YES**), and **QUEST**ions; unspecific semantics (**NIL**)). Furthermore, we use 8 abstract syntactic and 17 abstract semantic categories for the analysis at the phrase level. Since the flat syntactic and semantic analysis is learned, the approach does not crucially rely on these particular categories [24]. A more detailed description of these categories and flat syntactic and semantic analysis can be found in previous work [20, 23].

		Schön (Fine)	bis (until)	dann (then)	auf (good)	Wiedersehen (bye)
Basic	syntax	adjective	preposition	adverb	preposition	noun
	semantics	yes	here	time	nil	nil
Abstract	syntax	modus group	prepositional group		prepositional group	
	semantics	confirmation	time at		miscellaneous	

Fig. 2. Utterance with its flat representations. Note that the categories refer to the German utterance and cannot correspond correctly to the literal English word translation in all cases.

3 Architecture for speech/language integration

For fault-tolerant incremental speech processing and flat language analysis we developed the SCREEN system. The system currently contains six main parts. The parts are the *speech construction part* for constructing sentence hypotheses from word hypotheses delivered by a speech-recognizer, the *speech evaluation part* for computing syntactic and semantic plausibilities needed for choosing the best sentence hypotheses, the *category part* for flat syntactic and semantic analysis at the word and phrase level, the *correction part* for detecting and dealing with often occurring errors (repairs) in spontaneous language, the *subclause part* for detecting subclauses, and finally the *case frame part* for providing case frame representations of an utterance.

All parts consist of several modules which are either connectionist or symbolic. All modules communicate via a common message structure and allow parallel and incremental processing. The communication mechanism does not distinguish between symbolic and connectionist modules and is the basis for integration in our hybrid connectionist architecture SCREEN. The main goal in SCREEN is to push connectionist learning and a flat screening analysis as far as possible in order to provide fault-tolerant processing.

In previous work we gave a detailed description and analysis of the category part for flat syntactic [23] and semantic analysis [20] as well as the correction part [19] so that we will not go into details of the syntactic and semantic analysis here. In this article we will explain and discuss the new speech-related parts (speech construction part and speech evaluation part) within the general architecture.

3.1 Speech construction part

The speech construction part forms the interface between the speech recognizer[2] and the flat screening analysis[3]. Therefore the incremental stream of incoming

[2] Currently we use word hypotheses provided by our project partners at the University of Karlsruhe and the University of Hamburg.

[3] The speech recognizer and the speech construction part communicate via the INTARC/ICE communication environment [1].

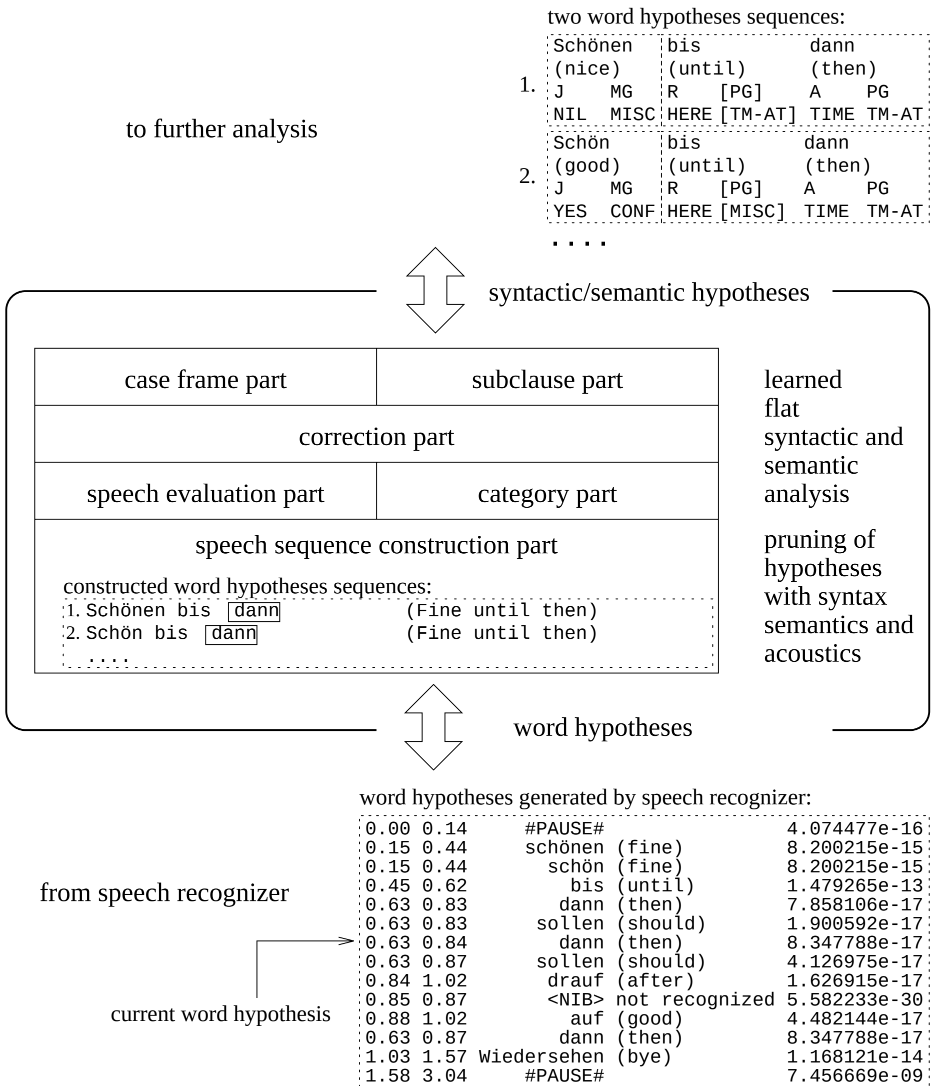

Fig. 3. Overview of the SCREEN system with example word hypotheses from tab.1.

word hypotheses is combined to an incremental stream of sentence hypotheses (see figure 3). These incremental partial sentence hypotheses become part of SCREEN's hypothesis space. The hypothesis space is pruned based on a combination of the acoustic plausibility (given by the speech recognizer), the syntactic plausibility, and the semantic plausibility (computed by modules of the speech evaluation part, see section 3.3). A sentence hypothesis may be dropped from the

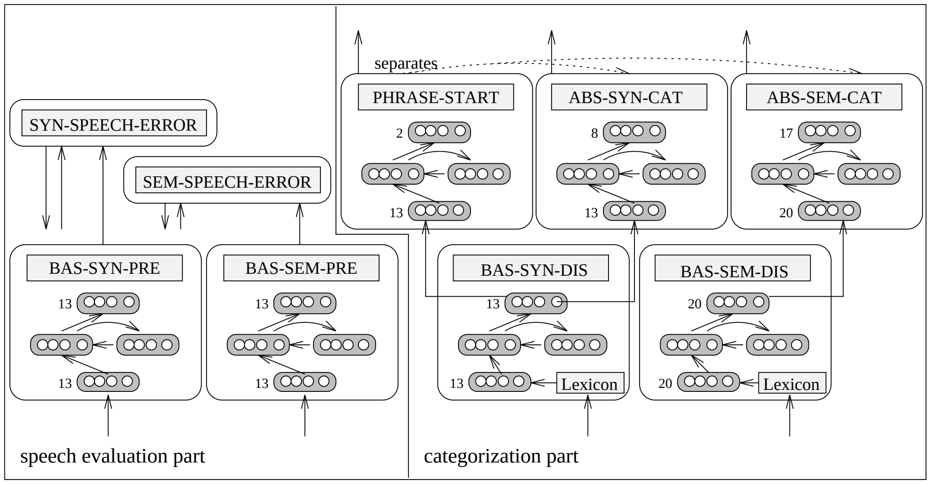

Fig. 4. Overview of the speech evaluation and categorization parts.

hypothesis space if its plausibility is too low or if it cannot be completed by any of the observed word hypotheses from the speech recognizer. The speech construction part combines the processing of individual sentence hypotheses with the parallel computation of competitive sequence hypotheses in the hypothesis space.

3.2 Categorization part

The computation of the syntactic and semantic plausibility of a partial sentence hypothesis depends on the basic syntactic disambiguation (BAS-SYN-DIS) and the basic semantic disambiguation (BAS-SEM-DIS). These disambiguations are also input for the abstract categorization modules which compute a phrase delimiter (PHRASE-START), abstract syntactic phrase categories (ABS-SYN-CAT), and abstract semantic phrase categories (ABS-SEM-CAT). Furthermore the output of the disambiguation is input to the prediction modules of the speech evaluation part.

The disambiguation in the modules BAS-SYN-DIS and BAS-SEM-DIS is based on simple recurrent networks and a syntactic and semantic lexicon (see figure 5). The current word of a sentence hypothesis is retrieved from the lexicon with its potentially ambiguous representation. The ambiguous representation is input to the network. The category of the unit with the highest activation in the output layer is chosen as the disambiguated category of the current word. The context layer from the previous disambiguations allows previous words to influence the current sentence hypothesis.

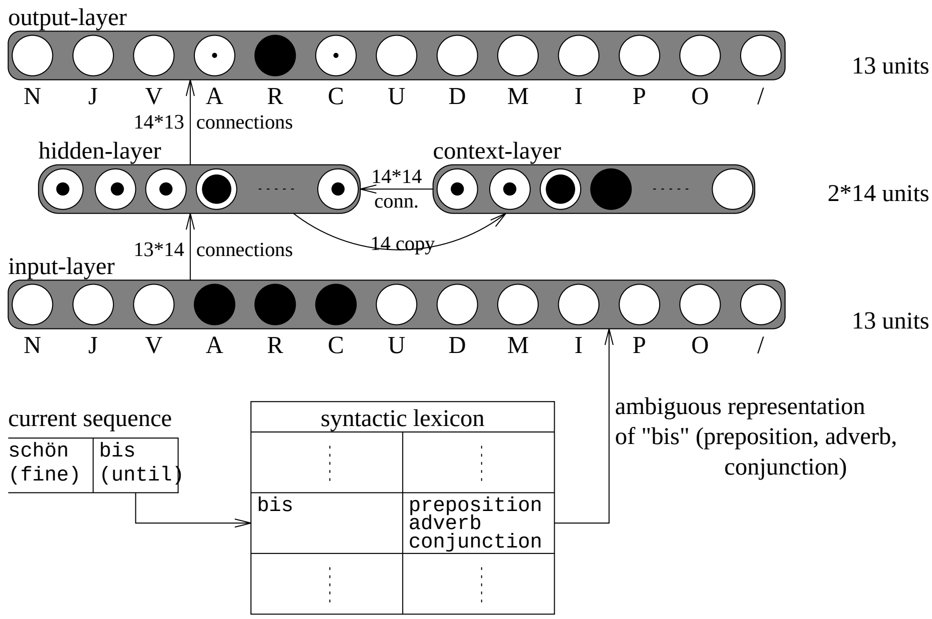

Fig. 5. Basic syntactic disambiguation for the sentence hypothesis "Schön bis ⋯" ("Fine until ⋯").

3.3 Speech evaluation part

The speech evaluation part computes the incremental syntactic and semantic plausibility of a sentence hypothesis. It uses the modules SYN-SPEECH-ERROR, SEM-SPEECH-ERROR, BAS-SYN-PRE, and BAS-SEM-PRE (see figure 6). The output of the BAS-SYN-PRE (resp. BAS-SEM-PRE) module is a vector of plausibilities for the prediction of categories. For instance the categories **N**oun, **D**elimiter, and nu**M**eral are very plausible to follow "bis" ("until") in figure 7 while **A**dverb, **I**nterjection, **P**ast participle, and **O**ther are less plausible but still possible. For computing a syntactic (resp. semantic) sentence plausibility based on possible category plausibilities we need to select a particular category. Since BAS-SYN-DIS (resp. BAS-SEM-DIS) disambiguates the categories we take the plausibility of the prediction vector for that category, which was found by the disambiguation network. This selection is done by SYN-SPEECH-ERROR for the syntactic sentence plausibility and by SEM-SPEECH-ERROR for the semantic sentence plausibility. If the syntactic category of the occurring current word is among the predicted syntactic categories at the directly preceding word, then this sentence hypothesis is considered plausible, otherwise implausible.

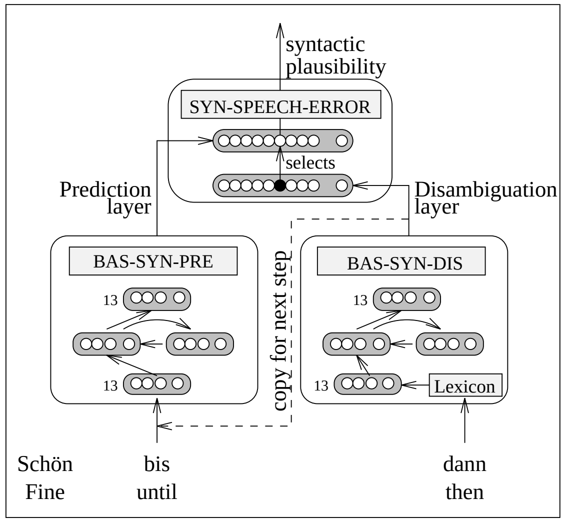

Fig. 6. Computing the syntactic plausibility for a sentence hypothesis.

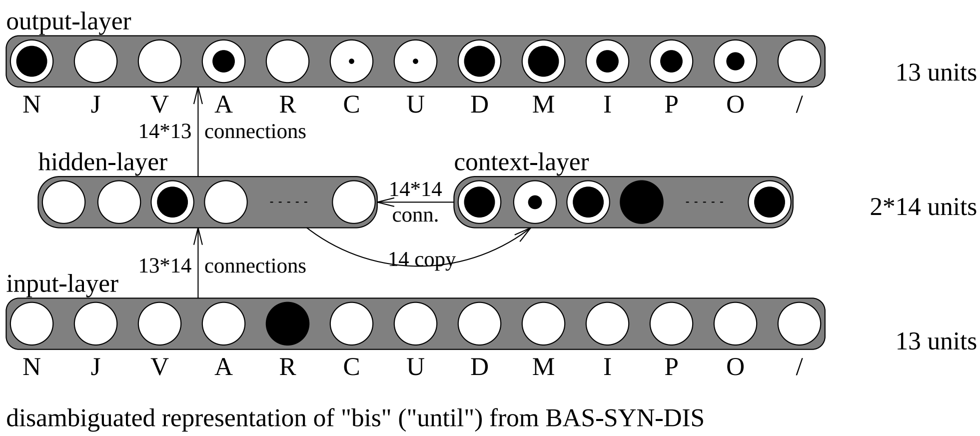

Fig. 7. Basic syntactic prediction for the sentence hypothesis "Schön bis $\cdots$" ("Fine until $\cdots$"). The basic syntactic categories were explained in section 2.2.

4 Improvement in the hypothesis space

In this section we illustrate the performance of learning a flat basic syntactic and semantic disambiguation and prediction. The overall performance for training and test sets is illustrated in table 2. The results for BAS-SYN-DIS, BAS-SEM-DIS, BAS-SYN-PRE, and BAS-SEM-PRE are based on a training set of 813 training words from 64 training sentences and 1543 test words from 120 unknown test sentences. We count a training or test instance as assigned correctly if the output unit with the maximal activation is equal to the desired category; otherwise we count it as an error. For the prediction networks we count if the desired category is within the categories with the highest five activations. This is necessary since words from more than one syntactic category can follow a particular word. Our results illustrate that 85%–97% of the training and 72%–89% of the test set have been learned. This performance allows us to use these networks in the speech evaluation part for the syntactic and semantic plausibility computation. Based on these combined acoustic, syntactic, and semantic plausibilities, first preliminary tests on 184 sentences show that the quality of the selected sentence hypotheses of the system could be increased by about 30% using acoustic and syntactic plausibilities and by about 50% using acoustic, syntactic, and semantic plausibilities.

Module	No. of units			correct assignments	
	I	H	O	train	test
BAS-SYN-DIS	13	14	13	97%	89%
BAS-SEM-DIS	20	14	20	95%	83%
BAS-SYN-PRE	13	14	13	89%	81%
BAS-SEM-PRE	20	14	20	85%	72%

Table 2. Training and generalization performance for disambiguation and prediction networks

5 The running system

In this section we describe the runtime behavior of SCREEN and we illustrate the flat analysis. Figure 8 gives an example of the SCREEN environment. Each horizontal stream represents a single sentence hypothesis from the hypothesis space. If there are more hypotheses in the hypothesis space than can be displayed they can be made visible by using the vertical scroll bar. The horizontal scroll bar is used for sentence hypotheses which are longer than the current display. The buttons allow the system to be stopped at a particular step for a deeper analysis of the system results.

Each word hypothesis of a sentence hypothesis is represented by a number of units. These units represent the flat syntactic and semantic analysis and the confidence for the sentence up to the current word. While the displayed units only show the highest activation for a categorization their contents can be inspected in more detail by a mouse click to a displayed unit. Then a zoom window occurs which displays all output units for a categorization (see bottom of figure 8).

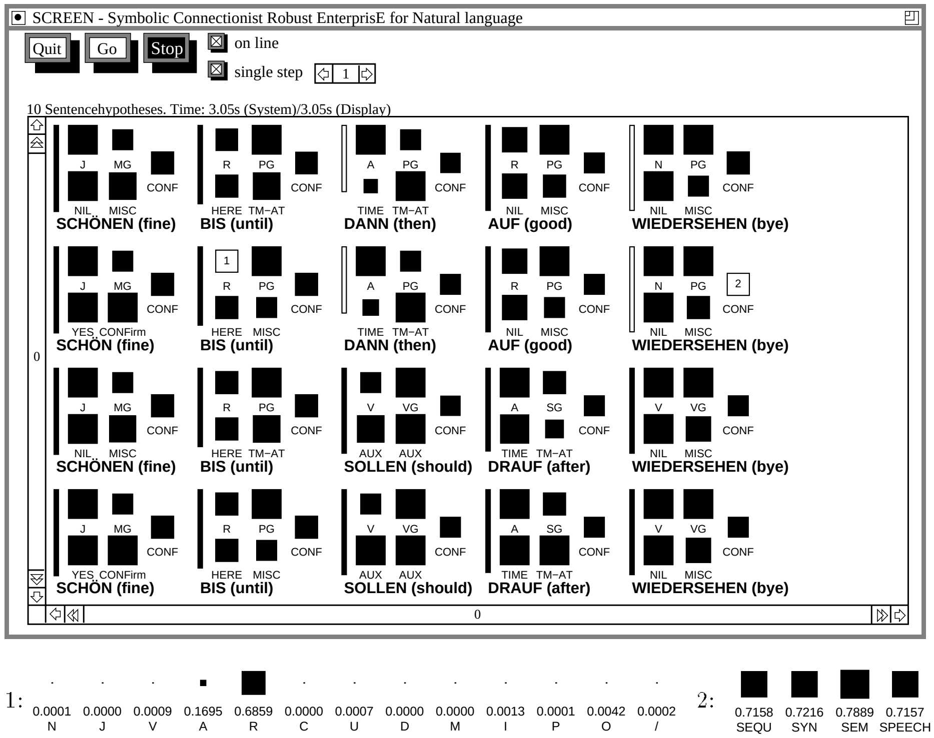

Fig. 8. Snapshot for sentence "Fine, until then, see you". Note that the categories refer to German rather than English. The abbreviations are: Ad**J**ective, p**R**eposition, **A**dverb, **N**oun, **V**erb; Modus **G**roup, **P**reposition **G**roup, **V**erb **G**roup, **S**pecial **G**roup; **NIL** or unspecific semantic, **HERE** time or location, **TIME** object, **YES** or agreement, **AUX**iliary statement; **MISC**ellaneous, **T**i**M**e-**AT**, **CONFIRM**ation. The white boxes in the second displayed sentence are zoomed and shown below the SCREEN window. The abbreviations for the right zoom window are: Overall plausibility for the **SEQU**ence up to the current word, **SYN**tactical resp. **SEM**antical plausibility, acoustic score from the **SPEECH**-recognizer.

For each word hypothesis there are five activations as square boxes and one as vertical bar. If a word starts a new phrase the vertical bar is filled; its length corresponds to the activation for the phrase-delimiter detector. The upper left box shows the basic syntactic category for the word, the lower left the basic semantic category. The upper right box shows the abstract syntactic category for a phrase; the lower right the abstract semantic category for a phrase. Based on previous experiments [21, 22] it is advisable to take the leftmost abstract syntactic category of a phrase as final syntactic interpretation for the whole phrase and the rightmost semantic abstract category of a phrase as final semantic interpretation of the phrase. The rightmost box shows the combined syntactic, semantic, and acoustic plausibility for the sentence up to the current word.

Figure 8 shows a part of the hypothesis space for the sentence "Fine, until then, good bye". The first hypothesis produced for the sentence "Schön, bis dann, auf Wiedersehen" ("Fine, until then, good bye") is "Schönen, bis dann, auf Wiedersehen" ("Fine, until then, good bye") which is syntactically and semantically correct although it contains a minor inflection incorrectness in German endings. Additional morphological knowledge could be used to avoid such cases. Now we will look at the second best proposed sentence hypothesis. The sentence starts with the word "Schön" ("Fine") which is identified as ad**J**ective and agreement (**YES**). The system interprets this word as a phrase which is a **CONFIRM**ation and a **M**odus **G**roup. **CONF**idence values for all words within the displayed hypothesis space are very high since implausible sentence hypotheses are dropped from the hypothesis space. The second word "bis" ("until") is a p**R**eposition and a time or location point (**HERE**). It introduces the new phrase "bis dann" ("until then"). The starting phrase is a **P**repositional **G**roup. The second word of this phrase is the **A**dverb "dann" ("then") which is a **TIME**-statement. The phrase determines a **T**i**M**e-**AT** which something happens. The final phrase is "auf Wiedersehen" ("good bye") which is a **P**repositional **G**roup introduced by the p**R**eposition "auf" ("good") and a **MISC**ellaneous-statement. The semantics for the word "auf" ("good") is unspecific (**NIL**). The final word "Wiedersehen" ("bye") is found to be a **N**oun also with unspecific semantics (**NIL**) in this domain since it belongs to a **MISC**ellaneous "politeness formula".

6 Discussion and related work

Learning: Our system differs from symbolic parsers, like for instance that of Mellish [11], and structured connectionist parsers, like that of Howells [6], in its ability to learn. As we have pointed out learning is essential for processing faulty real-world spontaneous language. Learning reduces knowledge engineering, avoids the design of a rule base and increases the portability of a system.

Fault-tolerance: Fault tolerance is extremely important for spoken language analysis because of spontaneity, irregularity and unforeseen input. Since the number of occurring errors and unforeseen constructions is unrestricted, a system must be able to deal with them without breaking. SCREEN possesses an explicit fault-tolerance represented by the modules of the correction part as well as an implicit fault-tolerance based on the inherent fault-tolerance and the similarity-based processing of connectionist networks [14].

Screening analysis with flat representations: We have developed and implemented a screening approach to speech/language analysis with a modular hybrid connectionist architecture. The system does a flat fault-tolerant syntactic analysis and semantic analysis of real-world spontaneous language. Within the screening approach we use a flat learned, fault-tolerant interpretation. The screening analysis is the basis for further analysis, for instance flat translation, functional grammar analysis as well as for the syntactic and semantic improvement of the hypothesis space.

The *related work* that is probably closest to ours is the PARSEC system [7]. PARSEC is a hybrid connectionist system and part of the speech translation system JANUS [17]. The aim of PARSEC is to generate case role representations based on sentences. It possesses several connectionist modules which trigger symbolic transformation rules. In contrast SCREEN offers also the opportunity of explicit fault tolerance in the correction part and SCREEN works on an incremental stream of real-world word hypotheses rather than manually transcribed sentences [7].

Hybrid Architecture: SCREEN is one of the first modular hybrid connectionist systems which integrate speech and language processing by using learned connectionist flat representations. We try to use connectionist networks wherever possible but also use symbolic processing wherever necessary. The system control structure and the message passing have been implemented symbolically as well as several small modules for simple equality tests. The connectionist modules have a symbolic interface and use a general message structure for communication. Therefore from the outside of the modules there is no difference between symbolic and connectionist modules. Modular hybrid architectures in other approaches have also been proven to be useful for common sense reasoning [14] and hybrid text processing [25].

Parallelism: Parallel processing is possible at different levels. First, there is the inherent parallelism of connectionist networks within the modules. Second, at the module level there are many modules running in parallel. Finally, sentence hypotheses in the hypothesis space can be processed in parallel.

Incrementality: Our approach is designed for incremental analysis. That is, we start the analysis with the first incoming word hypotheses from the speech recognizer. This allows us to integrate knowledge from the flat syntactic and semantic analysis for improving the hypothesis space. Therefore it is possible to reduce the hypothesis space at very early stages. The incremental analysis also allows SCREEN to start with further processing as early as possible.

7 Conclusions

We have described SCREEN, a hybrid connectionist system for incremental spoken language analysis. We have shown that flat connectionist analysis can be used to improve the choice of good sentence hypotheses. The system is based on the key principles of *learning*, *fault-tolerant processing*, and *flat screening* analysis. Based on the concepts in SCREEN we argue for a hybrid solution for the integration of speech and language processing. We use connectionist networks wherever possible and symbolic computation wherever necessary. The learned representation makes it easy to port the system to other corpora and domains. The fault-tolerance allows SCREEN to deal with irregularities generated by a human or a speech recognizer. Furthermore, the flat representations provide a new basis for speech/language processing, and support the acoustic, syntactic and semantic processing of real-world spoken utterances.

Acknowledgments

This research was funded by the German Research Association (DFG) under contract DFG Ha 1026/6-2 and by the German Federal Ministry for Research and Technology (BMBF) under Grant #01IV101A0. We would like to thank S. Haack, M. Löchel, M. Meurer and M. Schrattenholzer for their work on SCREEN. Special thanks to U. Sauerland for his work on the speech construction and evaluation parts.

References

1. J. W. Amtrup. ICE: INTARC communication environment. User guide and reference manual. VM-Techdok 14, Universität Hamburg, November 1994.
2. J. W. Amtrup, A. Hauenstein, C. Pyka, V. Weber, S. Wermter. An outline of the Verbmobil project with focus on the work at the University of Hamburg. In P. Mc Kevitt, ed., *Proceedings of the AAAI-94 Workshop on Integration of Natural Language and Speech Processing.* Seattle, Washington, USA, August 1994.
3. J. L. Elman. Finding structure in time. *Cognitive Science*, 14(2):179–211, 1990.
4. S. J. Hanson, J. Kegl. PARSNIP: A connectionist network that learns natural language grammar from exposure to natural language sentences. In *Proceedings of the 9th Annual Conference of the Cognitive Science Society*, pp. 106–119. Seattle, Washington, 1987.
5. A. Hauenstein, H. H. Weber. An investigation of tightly coupled time synchronous speech language interfaces using a unification grammar. In *Proceedings of the AAAI-94 Workshop on the Integration of Natural Language and Speech Processing.* Seattle, Washington, 1994.
6. T. Howells. VITAL: A connectionist parser. In *Proceedings of the 10th Meeting of the Cognitive Science Society*, pp. 18–25. Montreal, Canada, 1988.
7. A. N. Jain. Generalization performance in PARSEC - a structured connectionist parsing architecture. In J. E. Moody, S. J. Hanson, R. R. Lippmann, eds., *Advances in Neural Information Processing Systems 4*, pp. 209–216. Morgan Kaufmann, San Mateo, CA, 1992.
8. D. Jurafsky, C. Wooters, G. Tajchman, J. Segal, A. Stolcke, N. Morgan. Integrating experimental models of syntax, phonology, and accent/dialect in a speech recognizer. In *Proceedings of the AAAI-94 Workshop on the Integration of Natural Language and Speech Processing.* Seattle, Washington, July/August 1994.
9. J. L. McClelland, A. H. Kawamoto. Mechanisms of sentence processing: Assigning roles to constituents of sentences. In J. L. McClelland, D. E. Rumelhart, The PDP research group, eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 2., Psychological and Biological Models, chapter 19, pp. 272–331. MIT Press, Cambridge, MA, 1986.
10. C. McMillan, M. C. Mozer, P. Smolensky. Rule induction through integrated symbolic and subsymbolic processing. In J. E. Moody, S. J. Hanson, R. P. Lippmann, eds., *Advances in Neural Information Processing Systems 4*, pp. 969–976. Morgan Kaufmann, San Mateo, CA, 1992.
11. C. S. Mellish. Some chart-based techniques for parsing ill-formed input. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pp. 102–109, 1989.

12. R. Miikkulainen. *Subsymbolic Natural Language Processing. An integrated model of scripts, lexicon and memory.* MIT Press, Bradford Book, Cambridge, MA, 1993.
13. M. C. Mozer. Neural net architecture for temporal sequence processing. In A. Weigend, N. Gershenfeld, eds., *Predicting the future and understanding the past.* Addison-Wesley Publishing, Redwood City, CA, February 1993.
14. R. Sun. *Integrating Rules and Connectionism for Robust Common Sense Reasoning.* Wiley and Sons, New York, 1994.
15. G. Towell, J. W. Shavlik. Interpretation of artificial networks: Mapping knowledge-based neural networks into rules. In J. E. Moody, S. J. Hanson, R. R. Lippmann, eds., *Advances in Neural Information Processing Systems 4*, pp. 977–984. Morgan Kaufmann, San Mateo, CA, 1992.
16. W. von Hahn, C. Pyka. System architectures for speech understanding and language processing. In G. Heyer, H. Haugeneder, eds., *Appl.Ling.* Wiesbaden, 1992.
17. A. Waibel, A. N. Jain, A. McNair, J. Tebelskis, L. Osterholtz, H. Saito, O. Schmidbauer, T. Sloboda, M. Woszczyna. JANUS: Speech-to-speech translation using connectionist and non-connectionist techniques. In J. E. Moody, S. J. Hanson, R. R. Lippmann, eds., *Advances in Neural Information Processing Systems 4*, pp. 183–190. Morgan Kaufmann, San Mateo, CA, 1992.
18. N. Ward. An approach to tightly-coupled syntactic/semantic processing for speech understanding. In *Proceedings of the AAAI-94 Workshop on the Integration of Natural Language and Speech Processing.* Seattle, Washington, 1994.
19. V. Weber, S. Wermter. Artificial neural networks for repairing language. In *Proceedings of the 8^{th} International Conference on Neural Networks and their Applications.* Marseilles, FRA, December 1995.
20. V. Weber, S. Wermter. Towards learning semantics of spontaneous dialog utterances in a hybrid framework. In J. Hallam, ed., *Hybrid Problems, Hybrid Solutions — Proceedings of the 10^{th} Biennial Conference on AI and Cognitive Science*, pp. 229–238. Sheffield, UK, 1995.
21. S. Wermter, M. Löchel. Connectionist learning of flat syntactic analysis for speech/language systems. In M. Marinaro, P. G. Morasso, eds., *Proceedings of the International Conference on Artificial Neural Networks* , vol. 2, pp. 941–944. Sorrento, Italy, 1994.
22. S. Wermter, U. Peters. Learning incremental case assignment based on modular connectionist knowledge sources. In P. Werbos, H. Szu, B. Widrow, eds., *Proceedings of the World Congress on Neural Networks*, vol. 4, pp. 538–532. San Diego, CA, 1994.
23. S. Wermter, V. Weber. Learning fault-tolerant speech parsing with SCREEN. In *Proceedings of the 12^{th} National Conference on Artificial Intelligence*, vol. 1, pp. 670–675. Seattle, Washington, 1994.
24. S. Wermter, V. Weber. Artificial neural networks for automatic knowledge acquisition in multiple real–world language domains. In *Proceedings of the 8^{th} International Conference on Neural Networks and their Applications.* Marseilles, FRA, December 1995.
25. S. Wermter. *Hybrid Connectionist Natural Language Processing.* Chapman and Hall, London, UK, 1995.