LOGON

Technical Report # 2007-21

Extending LOGON:

Translating Other Language Pairs

(Draft of November 30, 2008)

Francis Bond

*National Institute of Information
and Communications Technology*

Contents

1 Introduction

This report outlines some of the current users of the LOGON architecture outside of the main Norwegian-English project. We discuss the general background (§ 2), then discuss the largest new system: Japanese↔English (§ 3). This is followed by a discussion of some issues that came up when adding Japanese↔English (§ 4). The report finishes with a brief list of other users of the system (§ 5).

2 Background

Semantic transfer is a general approach, and can be applied to any language pair. The LOGON architecture itself is language agnostic. In this chapter we describe in detail its use in building Japanese↔English systems, and mention some other uses.

The MRS framework is quite general and does not constrain the meaning representation very much. However, many grammars that use MRSshare a common core of assumptions laid out in the Grammar Matrix (§ 3.2: Bender, Flickinger, & Oepen (2002)). This backbone makes it easier to postprocess the semantics of different grammars — grammars compatible with the Matrix are already harmonized to some degree.

The first major extension of the LOGON system was to Japanese↔English. We took the Japanese grammar JACY (Siegel & Bender, 2002), originally developed in the Verb*mobil*project (Wahlster, 2000), along with the ERGand built two translation systems JaEn (Japanese-to-English) and EnJa (English-to-Japanese) (Bond, Oepen, Siegel, Copestake, & Flickinger, 2005). These systems take advantage of the existing translation rule types in the Norwegian-English system, sharing part of the type hierarchy.

Japanese and English are from completely different language families, and thus are fundamentally different at the syntactic level. One major difference is in word order — in Japanese, dependents normally precede their heads. For example, in a clause the verb is final (SOV): *[inu-ga]*$_{Subj}$ *[hone-o]*$_{Obj}$ *[tabeta]*$_{V}$ "the dog ate a bone (lit: dog bone ate)". Within a clause, the order of the dependents can vary quite freely. The relation between nouns and verbs is shown by postpositional particles (*joshi* in traditional Japanese grammar). Noun phrases are often omitted, particularly in speech, but also in formal edited text. Honorification can be marked on either nouns or verbs.

Another major difference is that Japanese does not mark number morphologically (singular/plural) or definiteness explicitly: there is no equivalent to the English articles (*a/the/*ϕ).

Typologically different languages are harder to translate between (see, e.g.

Koehn (2005)): Japanese↔English is therefore an even more challenging task than Norwegian↔English.

To test the applicability over a range of languages, we have also built proof-of-concept systems from Korean to Japanese, using the Korean Reference Grammar (Kim & Yang, 2003) and between Norwegian and Japanese using NorSource.

3 System Description: JaEn/Enja

The basic architecture of the Japanese↔English systems is the same as that of the LOGON Norwegian→English system. We use the same English grammar (the ERG: (Flickinger, 2000)) for English, and another HPSG-based grammar, Jacy (§ 3.1), for the Japanese. These are two of the grammars that have contributed to the multilingual matrix, and have been developed loosely in parallel. The grammars therefore share the same basic ideas about semantics. However, there has been far less work on explicit harmonization than for NorGram and the ERG.

We took advantage of the bi-directonal nature of the Japanese and English grammars to build both Japanese-to-English and English-to-Japanese systems. However, development to date has focused on Japanese-to-English. This technical report describes the system largely as it was in 2006, the system is further described in Bond et al. (2005), Nichols, Bond, Appling, & Matsumoto (2007)

3.1 Jacy

The Jacy grammar is an HPSG-based grammar of Japanese which originates from work done in the Verb*mobil* project (Siegel, 2000) on machine translation of spoken dialogues in the domain of travel planning. It has since been extended to accommodate written Japanese and new domains (such as automatic email response and parsing machine readable dictionaries).

The lexicon contains around 47,000 lexemes. The system also includes a mechanism to assume default lexical types for items that can be POS tagged by the ChaSen tokenizer and POS tagger (Asahara & Matsumoto, 2000), but are not included in the HPSG lexicon. As the grammar is developed for use in a variety of applications, it treats a wide range of basic constructions of Japanese. In the multilingual context in which this grammar has been developed, a high premium is placed on parallel and consistent semantic representations between grammars for different languages. Ensuring this parallelism enables the reuse of the same downstream technology, no matter which language is used as input.

3.2 The Grammar Matrix

Essential to the idea of developing machine translation systems applicable to multiple language pairs are the dual concerns of facilitating and harmonizing the construction of the grammars and lexicons. Facilitation addresses the concern that the construction of a large grammar and lexicon from scratch is an extremely time-consuming task; harmonization addresses the desirability of providing grammars and lexicons of different languages with a certain degree of uniformity, so as to enhance the cross-linguistic applicability of systems drawing on deep processing. To accommodate both of these concerns, a strategy of multilingual grammar engineering has been defined based on a sub-grammar called "The HPSG Grammar Matrix" (Bender et al., 2002). The Matrix consists of a skeleton of grammatical and lexical types, combined with a system of semantic representation—Minimal Recursion Semantics. It therefore constitutes a possible formal backbone for a large scale grammar of—in principle—any language. New grammar resources (e.g. for Italian and Norwegian) were built using the Matrix as a 'starter-kit for grammar writing'. Three existing grammars (English, German and Japanese) were adapted to the Matrix restrictions.

Using the Matrix simplifies the translation process by normalizing not just the type hierarchies but also the names used for common types. This is a prerequisite for the transfer Matrix — any types used in common transfer rule templates should be shared by all grammars. Examples of these universal types are *named_rel* for named entities, *pron_rel* for pronouns and *compound_rel* for noun-noun compounds.

In the current implementation there are some differences between the grammars for some of these predicates: these are trivially dealt with by accomodation rules: for example *place_rel* in the ERGis *place_n_rel* in Jacy. We chose one relation as canonical and map the other into it as the first step in one direction, and the last step in the other. These differences are not so important for the Norwegian↔English system, as rules were only being created for one language pair in one direction. As we smooth away these rough edges, less effort will be necessary to add a new Matrix compatible language pair.

From the point of view of machine translation, it would be helpful if the Matrix covered a wider range of phenomena. For example dates and times are often syntactically idiosyncratic, but share a relatively language independent semantics. A harmonized semantics would greatly simplify translation. Of course, it would still not be trivial. For example, while months and days are unambiguous in translation between Japanese and English, years are not: *3nen* "year 3" could be the third year of the current emperors reign (i.e., 1991) or 2003, neither of which would normally be written as just "3" in English.

3.3 System Development

Similarly to the Norwegian↔English system (Lønning et al., 2004) the creation of the transfer grammar comprised of three sub-tasks: (a) the identification and encoding of abstract transfer correspondence types, each mapping a piece of source language semantics to a corresponding target language structure—for example a Norwegian noun – noun compound into an adjective – noun configuration; (b) the instantiation of correspondence types with actual sets of semantic predicates, as associated to words and phrases from the LOGON vocabulary; and (c) the leveraging of a machine readable bi-lingual dictionary, complementing hand-built transfer instances with automatically created ones.

In order to increase coverage and make sure the system would scale, we chose to tightly integrate our transfer rule construction with an existing large-scale open source Japanese-English lexicon project: JMDict (Breen, 2004). We consider the transfer rules created from here the core rules, and we supplement them with hand-built rules. Our goal is to have a small set of hand crafted rules for closed class lexical items, such as determiners and prepositions, but to build open class rules from a bilingual dictionary and ultimately from aligned corpora.

3.4 Transfer Types

Most transfer phenomena were already covered by the correspondence types defined by Norwegian↔English. Only a handful of new rule-types have been necessary. One example was a rule to map zero-argument verbs to one-argument verbs with a subject *it rains* to *ame-ga furu* "rain falls". This was not necessary for the typologically close Norwegian-English pair, but is needed for Japanese↔English.

We often found it possible to reuse existing correspondence types by adding extra accommodation rules. For example, Japanese has a lexical class of verbal nouns, which can be used either with a support verb as predicates, or on their own as individuals. In Jaen, Japanese verbal nouns are analyzed as events, and they produce messages accordingly. When it is being used as a noun, *kenkyuu* "research" being kenkyuu_s_rel is wrapped with the relation noun-relation. We handle these constructions with a special rule that nominalizes the verbal noun by removing its event and the associated message, and replacing them with an entity when it appears as a noun:

```
vn-n_jf := monotonic_mtr &
[ CONTEXT.RELS < [ PRED "ja:udef_rel",
                   ARG0 #x0 ] >,
 IN [RELS <[PRED "ja:noun-relation",
         LBL #h6, ARG0 #x0, ARG1 #hp],
```

```
        [PRED "ja:proposition_m_rel",
         LBL #hp, ARG0 #ep, MARG #h5 ],
        [PRED #pred, LBL #h0, ARG0 #ep ]>,
     HCONS < qeq & [HARG #h5, LARG #h0 ]>],
 OUT [RELS <[PRED #pred, LBL #h6,
            ARG0 #x0 ]>,
      HCONS < > ] ].
```

In short, this rule type removes the noun-relation and all semantic relations resulting in the verbal noun's analysis as an event. This change makes it possible to treat verbal nouns identically to regular nouns in the rest of our transfer rules, eliminating the need to create multi-word transfer rules that have to distinguish between nouns and verbal nouns. This simplifies rule development significantly. And, thus, a rule to translate *kenkyuu* as the noun *research* can now be created using the standard noun template: –

```
kenkyuu_s-research_n-omtr := noun_mtr &
[IN.RELS <[PRED "_kenkyuu_s_rel"]>,
 OUT.RELS<[PRED "_research_n_1_rel"]>].
```

In addition, we can use the wide range of compound verb rules (n+n–n, adj+n-na and so forth) with both normal nouns and verbal nouns.

Another approach would have been to add further subtypes of all these rules (n+vn–n, vn+n–n, vn+vn–n, . . .), so far this has not proved necessary.

3.5 Transfer Instances (Automatic)

Nygaard, Lønning, Nordgård, & Oepen (2006) demonstrated that it is possible to learn transfer rules for some open category lexical items using a bilingual Norwegian→English dictionary. They succeeded in acquiring over 6,000 rules for adjectives, nouns, and various combinations thereof. Their method entailed looking up the semantic relations corresponding to words in a translation pair, and matching the results using simple pattern matching to identify compatible rule types. Our approach is an effort to generalize this approach by using rule templates to generate transfer rules from input source and target MRSstructures. it template mappings are used to identify translation pairs where there is a compatible rule type that can be used to create a transfer rule. A template mapping is a tuple consisting of: (i) a list of HPSG syntactic categories corresponding to the words in the source translation; (ii) a list of HPSG syntactic categories for the target translation words; and (iii) the name of the rule template that can be used to construct a transfer rule. For example, the template mapping ([noun], [adjective, noun], n-an) identifies a template that creates a rule to translate a Japanese noun into an English adjective-noun sequence. We use JMDict, the Japanese→Multilingual dictionary created by Jim Breen (Breen, 2004) to automatically acquire transfer rules. JMDict has approximately 110,000 main entries, with an additional 12,000 entries for computing and communications technology, and dictionary of over 350,000 proper names. Transfer rule generation is carried out in the following manner:

1. Look up all words in source language translation in HPSG grammar
   - Retrieve syntactic categories and MRSrelations
   - Enumerate every possible combination for words with multiple entries
   - Refactor results into separate lists of syntactic categories and MRSrelations
2. Repeat 1. for all words in target language translation
3. Map template mappings onto source and target syntactic categories
   - Translations that match indicate existence of compatible rule template
4. Create a transfer rule by calling the rule template with lists of MRSrelations as its arguments

The results of open category transfer rule acquisition from EDICT are summarized in Table 1. We have also extracted several thousand rules consisting of multiple words in length.

In addition to creating rules from a machine-readable dictionary, we also run the same machinery over aligned chunks from the phrase table of a statistical machine translation system. For this we used the open source sytem MOSES (Koehn et al., 2007). This adds anohter few thousand rules including words such as *suteki* "nice", where the dictionary only gave us "great" and "beautiful" as well as names such as *ganji-* "Gandhi".

3.6 Enhancing the Bilingual Dictionary

We have enhanced the JMdict lexicon in two ways. The first is an explicit distinction between transfer equivalents and explanations:

(1) *ten*

```
<gloss g_type="equ">spot</gloss>
<gloss g_type="exp">counter for goods or items</gloss>
```

The second is to explictly separate disjunctive entries:

11751	noun_mtr
7888	n+n_n+n_mtr
5638	n+n_adj+n_mtr
5575	n+n_n_mtr
2325	n_n+n_mtr
2294	n_adj+n_mtr
2101	noun_omtr
2038	intersective_attribute_mtr
1940	arg12_v_mtr
855	arg1_v_mtr
701	adj+n_adj+n_mtr
294	intersective_attribute_omtr
194	arg12_v_omtr
165	v_adj_mtr
138	n_n+n_omtr
136	n_adj+n_omtr
116	arg1_v_omtr
80	adj_v_mtr
22	adj_v_omtr
20	n+n_adj+n_omtr
17	n+n_n_omtr
14	v_adj_omtr
10	n+n_n+n_omtr

Table 1: Automatically Created Rule Types

(2) *denchi* /(n) farmland/rice field or paddy/

⇒ /(n) farmland/ rice field/ rice paddy/

These two extensions make it possible to produce transfer rules only for those entries which are true translations.

3.7 Transfer Instances (Manual)

In order to decide which semantic relations to write transfer rules for by hand, we used the automatically acquired translation rules in the above section and attempted to translate sentences from the BTEC* corpus. Whenever a relation failed to transfer, the system would be unable to generate a translation, and an error message was produced. We counted the relations and identified the most frequently occurring closed class relations as candidates for handcrafting a transfer rule. There are currently around 500 handcrafted rules in our system. A list of the 10 most common un-translatable relations and glosses of the translations we created are given in Table 2.

In handcrafting transfer rules for our system, we also encountered several linguistic problems that needed to be solved in order to achieve high-quality translation results, the most interesting of which was pronoun generation in English. Since our Japanese semantic analyses indicate when arguments of a predicate have been omitted, we came up with a small set of rules that checks what restrictions, if any are placed on the omitted arguments, and we replace them with underspecified English pronouns, since the nature of the omitted argument is unknown. This causes our system to generate "I/you/we/he/she/it/they" for every pronoun inserted, so to avoid an explosion in the number of translations, we only allow pronouns to be inserted for the first two argument slots (roughly corresponding to subject and objects). Other advances include the treatment of common modal verbs, and natural generation of determiners for negative clauses. We spent approximately one month on handcrafting transfer rules.

As development and testing data, we are currently using the Tanaka Corpus (Tanaka, 2001) and the ATR Basic Travel Expression Corpus as made available in the IWSLT 2005 evaluation campaign (Eck & Hori, 2005). As is indicated in its name, the BTEC* corpus consists of short spoken sentences taken from the travel domain. We selected it because is it a commonly used development set, making our results immediately comparable to a number of different systems, and because our Japanese HPSG parser can successfully analyze approx. 80% of its sentences, providing us with a good base for development. The BTEC* data supplied in the ITWSLT 2006 evaluation campaign consists of almost 40,000 aligned sentence pairs. Sentences average 10.0 words in length for Japanese and 9.2 words in length

Frequency	Semantic relation	Translation
25,927	"_ni_p_rel"	ni → in, to, into
25,056	"cop_id_rel"	da, desu → to be
22,976	"_no_p_rel"	X-no-Y → X Y, X's Y, Y of X
10,375	"_de_p_rel"	de → in, on, at, with
9,696	"rareru_rel"	-rareru → passive
9,528	"neg_v_rel"	-nai → negation
8,848	"_exist_v_rel"	aru → to be, to have
7,627	"_kono_q_rel"	kono → this
4,173	"tai_rel"	-tai → to want to
3,588	"_hour_n_rel"	-ji, toki → time, hour

Table 2: Most frequently occuring source language relations and their hand-crafted translations

for English. There are 11,407 unique Japanese tokens and 7,225 unique English tokens.

4 Some issues

4.1 Japanese/English Linguistic differences

Several phenomena have proved to be more of an issue when translating between English and Japanese than between Norwegian and English. These include omitted arguments (Japanese normally omits recoverable referents, rather than using pronouns), handling honorification (which is lexicalized in Japanese) and transliteration (it is ok to just output an unknown Norwegian word as is in NoEn, this is less acceptable in JaEn, because Japanese uses a completely different writing system(Nichols, Bond, & Matsumoto, 2007).

From the beginning, the JaEn project decided to automatically produce as many simple transfer rules as possible from the existing Japanese-English dictionary project JMDict (Breen, 2004; Bond & Breen, 2007). This involved both automatic extraction and enhancing the JMDict representation. We are now working on extracting rules from corpora.

We did a little harmonization to make life simpler. We changed the grammar encoding to utf-8. We have harmonized the treatment of titles, demonstratives and time expressions.

4.2 Multilingual MT and the Transfer Matrix

4.2.1 Predicate Name Spaces

We had some problems with rule translatings Japanese to English, and the resulting English then being treated as Japanese and translated again! For example *ame* → *rain* → *line*.

```
ame_n-rain_n-omtr := noun_mtr &
[IN.RELS <[PRED "_ame_n_1_rel"]>,
 OUT.RELS<[PRED "_rain_n_1_rel"]>].

rain_n-line_n-omtr := noun_mtr &
[IN.RELS <[PRED "_rain_n_1_rel"]>,
 OUT.RELS<[PRED "_line_n_1_rel"]>].
```

We deal with this by seperating the name spaces, in our case by simply prepending “ja:” to all Japanese predicates in the input.

4.2.2 Smuggling types into the MRS

Many of the uses of regular expressions in the transfer rules are effectively referring to types for strings.

For example, in the plural marker rule shown below, an NP marked in Japanese with a plural marker is given plural number, and the elementary predicate of the marker discarded:

```
plural-marker-jf := monotonic_mtr &
 [ CONTEXT [RELS < [ ARG0 #x0, RSTR #hr ] >,
    HCONs < qeq & [ HARG #hr, LARG #h0 ] > ],
   INPUT.RELS < [ PRED #pred, LBL #h0,
  ARG0 #x0 & [ PERS #pers, NUM #num ]],
[ PRED "~ja:.*_pl_rel",
  LBL #h1, ARG0 #e1, ARG1 #x0 ] >,
   OUTPUT.RELS < [PRED #pred, LBL #h0,
  ARG0 #x0 & [ PERS #pers, NUM pl ]] > ].
```

(3) *inu tachi* “dog + others” → *inu* “dogs”

There are four possible plural markers differing only in the level of respect they imply (*ra, domo, tachi, kata* from least to most polite). Jacy chose to mark them as strings, not types, but the transfer rule should match any of them. In this case we use a regular expression to, in effect, create a type on the fly, and match on that.

Similarly, there are rules that match on all verbs (_v_) or prepositions (_p_). As the transfer machinery is designed to work on types, it seems a pity not to use them here.

4.2.3 Bidirectional Rules

Many simple rules are effectively bidirectional, for example: *inu_n_1*x ↔ *dog_n_1*x, or even more complicated rules like *kiiro_n_1*x+*no_p*e,x,y ↔ *yellow_a_1*e,y. In principal, anything with no filter, context or regular expression in the rule could be reversed.

However, problems arise because the rules are ordered — if there are two translations for *inu* (*dog* and *spy*), then the first must be optional and only the final one obligatory. These ordering issues are not symmetrical: *spy* in turn has at least two translations *supai* and *inu*. At present, therefore, we take advantage of the fact that most of our rules are automatically generated from a bilingual transfer lexicon, and generate both directions separately. Hand-written rules must currently be written twice, once for Japanese→English and once for English→Japanese.

To minimize the amount of rewriting, we have added new attributes JA and EN, which map to INPUT and OUTPUT for Japanese→English and OUTPUT and INPUT for English→Japanese. This allows us to cut and paste a single rule for both directions.

5 Other extensions of the LOGON architecture

5.1 Further Prototypes

To encourage joint research with other sites, we also produced small prototype grammars for Norwegian↔Japanese, Korean→Japanese and Spanish→Japanese. The Norwegian↔Japanese system was used to test using enriched preposition semantics in translation.

5.2 EBMT: RBMT with learned rules

A completely different approach was take in a prototype German↔English system, which learns almost all of its rules from parallel text (Jellinghaus, 2007). This has the advantage that it is almost completely automatic, although initially only for those phenomena which align reasonably well.

5.3 Teaching

Finally, the MT system has been used in at least two university courses. In one, on grammar engineering at the University of Washington, it is used in the final session to translate between all the grammars constructed in the course. In another, on natural language processing at Nara Women's University, it was used to highlight problems in natural language processing, and show how solutions can be developed.

References

Asahara, M., & Matsumoto, Y. (2000). Extended models and tools for high-performance part-of-speech tagger. In *Proceedings of the 18th International Conference on Computational Linguistics* (pp. 21 – 27). Saarbrücken, Germany.

Bender, E. M., Flickinger, D., & Oepen, S. (2002). The grammar Matrix. An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammar. In *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics.* Taipei, Taiwan.

Bond, F., & Breen, J. (2007). Semi-automatic refinement of the JMdict/EDICT Japanese-English dictionary. In *13th annual meeting of the association for natural language processing* (p. 364-367). Kyoto.

Bond, F., Oepen, S., Siegel, M., Copestake, A., & Flickinger, D. (2005). Open source machine translation with DELPH-IN. In *Proceedings of the Open-Source Machine Translation workshop at the 10th Machine Translation Summit* (pp. 15 – 22). Phuket, Thailand.

Breen, J. W. (2004). JMDict: a Japanese-multilingual dictionary. In *Coling 2004 workshop on multilingual linguistic resources* (pp. 71–78). Geneva.

Eck, M., & Hori, C. (2005). Overview of the IWSLT 2005 evaluation campaign. In *Proc. of the international workshop on spoken language translation* (p. 11-32). Pittsburgh, USA.

Flickinger, D. (2000). On building a more efficient grammar by exploiting types. *Natural Language Engineering, 6 (1),* 15 – 28.

Jellinghaus, M. (2007). *Automatic acquisition of semantic transfer rules for machine translation.* Unpublished master's thesis, Universität des Saarlandes.

Kim, J.-B., & Yang, J. (2003). Korean phrase structure grammar and its implementations into the LKB system. In D. H. Ji & K. T. Lua (Eds.), *Proceedings of the 17th asia pacific conference* (p. 88-97). COLIPS Publications.

Koehn, P. (2005). Translating europarl. In *Mt summit x.* Phuket.

Koehn, P., Shen, W., Federico, M., Bertoldi, N., Callison-Burch, C., Cowan, B., Dyer, C., Hoang, H., Bojar, O., Zens, R., Constantin, A., Herbst, E., Moran, C., & Birch, A. (2007). Moses: Open source toolkit for statistical machine translation. In *Proceedings of the acl 2007 interactive presentation sessions.* Prague.

Lønning, J. T., Oepen, S., Beermann, D., Hellan, L., Carroll, J., Dyvik, H., Flickinger, D., Johannessen, J. B., Meurer, P., Nordgård, T., Rosén, V., & Velldal, E. (2004). LOGON. A Norwegian MT effort. In *Proceedings of the Workshop in Recent Advances in Scandinavian Machine Translation.* Uppsala, Sweden.

Nichols, E., Bond, F., Appling, D. S., & Matsumoto, Y. (2007). Combining resources for open source machine translation. In *Proceedings of the 10th International Conference on Theoretical and Methodological Issues in Machine Translation.* Skövde, Sweden.

Nichols, E., Bond, F., & Matsumoto, Y. (2007). Automatic transfer rule acquisition for semantic transfer based MT. In *Ipsg-sig-nl 178* (pp. 77–84). Nagoya.

Nygaard, L., Lønning, J. T., Nordgård, T., & Oepen, S. (2006). Using a bi-lingual dictionary in lexical transfer. In *Proceedings of the 11th conference of the European Association for Machine Translation.* Oslo, Norway.

Siegel, M. (2000). HPSG analysis of Japanese. In W. Wahlster (Ed.), *Verbmobil. Foundations of speech-to-speech translation* (Artificial Intelligence ed., pp. 265 – 280). Berlin, Germany: Springer.

Siegel, M., & Bender, E. M. (2002). Efficient deep processing of Japanese. In *Proceedings of the 3rd workshop on Asian language resources and international standardization at the 19th international conference on computational linguistics.* Taipei.

Tanaka, Y. (2001). Compilation of a multilingual parallel corpus. In *Proceedings of pacling 2001* (pp. 265–268). Kyushu. ((`http://www.colips.org/afnlp/archives/pacling2001/pdf/tanaka.pdf`))

Wahlster, W. (Ed.). (2000). *Verbmobil. Foundations of speech-to-speech translation.* Berlin, Germany: Springer.