Issues in Korean/English Machine Translation

Chung-hye Han, Meesook Kim, Martha Palmer

University of Pennsylvania

Feb. 10, 2000

Word Order and Scrambling

source	추가 공급물을 103 전위지원대대에게 사령부가 주었다. ({Chu-Ka} {Kong-Keup-Mul-eul} 103 {Ceon-wi-Ci-weon-Tae-Tae-e-Ke} {Sa-Ryeong-Pu-Ka} {Cu-eoss-Ta}.)
Glosser	additional supply_2 103 FSB_1 headquarter_0 gave
OTS MT system	Additional supply_2 103 FSB_1 headquarters_0 which you bite $\boxed{\text{gave}}$.
target	Headquarters_0 $\boxed{\text{gave}}$ 103rd FSB_1 additional supplies_2.
Penn/CGT	Headquarters_0 gave an additional supply_2 to a 103 forward support battalion_1.

Wh-question

source	무전기용 축전지는 보통 얼마나 오래 가는가? {Mu-Ceon-Ki-yong} {Chuk-Ceon-Ci-Neun} {Po-Thong} {eol-Ma-Na} {o-Rae} {Ka-Neun-Ka}?
Glosser	radio battery-Top usually how long last-Int
target	How long does a radio battery usually last?
Penn/CGT	How long does a radio battery usually last?

Empty Arguments

source	우리의 최근 요청을 수신했는가? ({u-Ri-yi} {Choe-Keun} {yo-Cheong-eul} {Su-Sin-Haess-Neun-Ka}?)
Glosser	our current request received
OTS MT system	$\boxed{\text{..... To}}$ receive an our recent request?
target	$\boxed{\text{Did (NP)}}$ receive our current request?
Penn/CGT	**Did you** receive our current request?

Morphological Analysis and Empty Arguments

source	IBP 화일을 검색하지 못했다면 지금 다시 보내겠다. (IBP {Hwa-il-eul} {Keom-Saek-Ha-Ci} {Mos-Haess-Ta-Myeon} {Ci-Keum} {Ta-Si} {Po-Nae-Kess-Ta}.)
Glosser	IBP file retrieve could not now again will send
OTS MT system	The IBP the file retrieve the cotton now not to be able you will send again.
target	If (NP1) could not retrieve IBP file, (NP2) will send (NP3) again now.
Penn/CGT	**If one** can not retrieve an IBP file, **one** will send **it** again now.

Lexical Ambiguity

- Postposition: (**{u-Ro}**) ⇔ as, by, to, with

source	보고서는 서면-**으로** 제출되었다. ({Po-Ko-Seo-Neun} {Seo-Myeon-**u-Ro**} {Ce-Chul-Toe-eoss-Ta}.)
gloss	report-Top document-as submitted
target	The report was submitted **as** a written document.
OTS MT system	*The report presented **with** the written document.
Penn/CGT	A report was submited **as** a written document.

source	부대는 전방-**으로** 이동했다. ({Pu-Tae-Neun} {Ceon-Pang-**u-Ro**} {i-Tong-Haess-Ta}.)
gloss	unit front-line-to moved
target	The unit moved **to** the front line.
OTS MT system	*The squad moved **with** the front line.
Penn/CGT	A unit moved **to** a front line.

- Verb: **{Ka-Ta}** ⇔ go, last, light verb

source	그 다음에 산기슭을 따라 동쪽으로 약 1 킬로 가게 되지요. ({Keu} {Ta-eum-e} {San-Ki-Seulk-eul} {Tta-Ra} {Tong-Ccok-eu-Ro} {yak} 1 {Khil-Ro} {Ka-Ke} {Toe-Ci-yo}.)
gloss	then mountain-base along eastward-to about 1 kilometer go
target	Then we would go east along the base of the hill for about a kilometer.
Penn/CGT	Then one goes about 1 Kilo to the East along a mountain base.
source	무전기용 축전지는 보통 얼마나 오래 가는가? ({Mu-Ceon-Ki-yong} {Chuk-Ceon-Ci-Neun} {Po-Thong} {eol-Ma-Na} {o-Rae} {Ka-Neun-Ka}?)
gloss	radio battery-Top usually how long last-Int
target	How long does a radio battery usually last?
Penn/CGT	How long does a radio battery usually last?
source	엔진이 장시간 과열하면 무리가 간다. ({en-Cin-i} {Cang-Si-Kan} {Kwa-yeol-Ha-Myeon} {Mu-Ri-Ka} {Kan-Ta}.)
gloss	engine-Nom long-period-of-time overheat-if overstrain-Nom cause-Decl
target	Excessive overheating over a long period of time will overstrain the engine.
Penn/CGT	When an engine is overheated a long time, it overstrains.

Modal Auxiliary, Empty Argument, Postposition

source	전투시에는 통신망을 잘 운용해야 합니다. (Ceon-Thu-Si-e-Neun Thong-Sin-Mang-eul Cal un-yong-Hae-ya Hap-Ni-Ta.)
gloss	combat-in communication-network well operate must
Penn/CGT	**One must** operate a communications network well **in** combat.