

번역품질평가의 현황과 과제

이 향
(고려대)

1. 들어가는 말

과거의 번역담론이 ‘어떻게 번역해야 하는가?’, 혹은 ‘좋은 번역이란 무엇인가?’ 등의 물음에 대하여 경험이나 직관을 토대로 한 조언이나 처방을 제시하는 수준에 머물렀다면, 번역학이 독립적 학문영역을 구축하기 시작한 1970년대 이후부터 최근까지의 번역학은 번역현상을 관찰, 기술하고 그로부터 번역행위를 지배하는 일반 원칙을 도출하는 것을 목적으로 하는 경험과학으로서의 면모를 강화하는 일에 주력해 온 것이 사실이다. 체스터만(Chesterman 201)은 이렇듯 번역학이 과거의 처방적(prescriptive) 담론에서 기술적(descriptive) 담론으로 옮겨간 것을 번역학 내부에서 일어난 가장 의미심장한 변화 중 하나로 꼽는다.

번역 품질에 관한 이론적 담론의 장은 어찌 보면 이러한 처방론과 기술론이 가장 치열하게 부딪히는 영역이기도 하다. 번역물에 대한 ‘평가’라는 것은 그것이 어떤 방식으로 이루어지든 일정 정도의 자의성과 ‘처방적’ 성격의 담론

을 필연적으로 내포한다. 무엇이 ‘좋은’ 번역이고 무엇이 ‘나쁜’ 번역인지를 평가한다는 것은 어떤 ‘특정한’ 방식으로 이루어진 번역이 좋은 번역이라는 평가자의 가치판단을 전제하므로, 당연히 번역이 어떤 특정한 방식으로 이루어질 것을 간접적으로 ‘처방’하는 결과를 낳기 때문이다. 모든 번역품질평가에 내포된 이러한 ‘처방적’ 성격은 번역평가에 대한 담론 전체를 주관적, 자의적인 것으로 오해하게 만들기도 하며, 이에 대한 반발로 번역에 대한 평가를 최대한 계량화, 객관화해야 한다는 또 다른 주장들을 낳게 된다. 그러나 다른 한편으로 평가에 대한 입장의 차이를 넘어서서, 번역품질의 관리와 평가가 실무적 필요이자 당위이라는 점에 대해서는 대체로 모두가 동의한다. 번역물에 대한 품질평가가 완전한 객관성을 담보하는 방식으로 이루어지지 않는 못하더라도, 자의성을 최소화하고 보다 효율적이고 체계적인 방식으로 수행될 수 있도록 해줄 해결책이 필요한 것이 사실이다.

현재 국내에서 번역 품질에 관한 담론, 보다 넓게는 ‘무엇이 좋은 번역인가’에 관한 담론은 양 갈래로 나뉘어 진행되고 있는 듯이 보인다. 그 첫 번째는 고전작품이나 주요 문학작품의 번역품질에 대한 논의를 핵심으로 하는 소위 ‘번역비평’ 담론이다. 사실 이러한 논의는 최근에 시작된 것은 아니며, 특정 작품의 번역을 서평(혹은 번역평)의 형식으로 지적하거나, 보다 포괄적으로는 우리의 열악한 번역 현실을 개탄하는, 더 나아가 졸속 번역으로 인한 인문학의 ‘위기’를 논하는 류의 글들을 우리는 수시로 접해왔다. 다만 최근에 와서 이러한 작업들이 보다 체계적이고 대대적인 방식으로 시도되고 있다는 점은 눈여겨 볼 필요가 있다. 대표적인 예로 영미문학연구회(이하 영미연) 번역평가사업단이 진행한 <영미고전문학 번역평가사업>을 들 수 있다. 총 4년에 걸쳐 36편 900여종에 달하는 영미문학 작품의 완역본들을 대상으로 질적 평가를 수행하고 신뢰할 만한 번역본(추천본)을 가려낸 영미연의 작업은 그간 ‘번역의 수준과 풍토’를 닦아내면서도 ‘종합적이고 실증적인 자료’는 찾아보기 힘들었던 국내 현실에서 의미 있는 결실이라고 할 수 있다(영미연 2005: 5). 현재 고려대학교 프랑스로프 작소설 번역평가 연구단에서 수행하고 있는 사업 역시 이러한 범주에 속한다고 볼 수 있다. 이상의 논의들은 주로 문학번역, 혹은 작품번역의 범주 내에서 기존 번역물들의 품질에 대한 대대적인 진단을 내리는 동시에 무엇이 좋은 번역인가에 대한 진지한 논의의 토대를 마련해 가고 있다.

국내 번역평가 담론의 다른 한 축을 구성하는 나머지 한 갈래는 소위 ‘번역 품질평가(TQA, Translation Quality Assessment)’라는 이름으로 번역학 내부에서 진행되고 있는 논의들이다. 번역교육의 과정에서 이루어지는 번역학습자의 번역능력평가, 특정 언어조합의 번역상에서 나타나는 문제점이나 오류 분석, 순차통역 및 동시통역의 품질평가 등을 포함하는 이러한 연구들은 원칙적으로 문학작품번역에서의 품질평가 문제를 배제하지는 않으나 주로 통번역의 교육과정에서 혹은 실무 번역 현장에서 발생한 구체적 수요로부터 출발한 연구들이라는 점에서 다소 차이가 있다.

의아한 것은, 이 두 갈래의 담론이 결국은 ‘번역의 품질’이라는 상위주제를 공유하고 있음에도 불구하고 서로 충분히 소통하지 못하고 있다는 점이다. 예를 들어 문학작품의 품질에 관한 기존 논의들은 번역학 내부에서 진행되어 온 TQA 담론을 참조하는 경우가 드물다. 영미언에서 수행한 번역비평 작업을 담은 총 1000여 페이지 상당의 『영미명작 좋은 번역을 찾아서 1, 2』는 번역의 질을 판별하는 기본범주로 ‘충실성’과 ‘가독성’을 선택하였다고 밝히고 각각의 개념에 대해 간략하게 설명하고 있을 뿐, 번역학 내부에서 진행되어 온 번역품질 평가기준에 대한 논의들은 전혀 언급하지 않는다(영미언 2005: 21). 이는 역으로도 마찬가지로 번역학 내부에서 번역평가에 대한 실무적 경험들을 토대로 개선되어 온 담론들은 문학번역 영역에서 꾸준히 누적되어 온 성찰이나 분석들에 충분한 주의를 기울이지 않는다. 물론 번역학 내부에서의 이러한 ‘파편화(fragmentation)’ 현상은 단순히 품질관련 논의에서만 발견되는 것은 아니며, 어찌 보면 번역에 관한 담론 전체를 특징짓는 현상이기도 하다. 이미 10여 년 전, 베이커(Baker 9)는 무수한 접근(approaches), 지평(horizons), 방법론들(methodologies)의 ‘급격한 변화’와 ‘파편화’라는 두 개의 키워드로 번역학의 현황을 요약하였다. 그리고 오늘날에도 이러한 변화와 파편화는 계속되는 것처럼 보인다. 문학번역과 실용번역 사이에는 거대한 벽이 존재하고 번역이론가들은 종종 실무와 단절되어 있으며 실무번역가들은 자신의 경험을 이론화하는 작업에 그다지 적극적이지 못하다. 번역 이론가들 내부에서도 언어학을 뿌리로 하는 연구자들은 문학으로부터 출발하여 번역을 설명하는 또 다른 그룹의 이론가들과 거의 소통불능의 상태에 있다. 이들은 ‘번역’이라는 공통의 주제를 논하면서도 각자 개별적 공간 안에서 개별적 담론을 생산할 뿐 서로 교류하지 않는다.

필자 역시 이러한 상황으로부터 자유롭지 못하며, 따라서 감히 위의 두 갈래를 아우르는 방식으로 번역평가를 논할 만한 역량은 없다. 본고는 위의 두 갈래 중 후자, 즉 번역학 내부에서 진행되고 있는 담론들을 개괄하되, 이를 바탕으로 다양한 분야에서 이루어지는 번역품질평가에 포괄적으로 적용 가능한 몇 가지 시사점들을 도출해 내는 선에서 만족하고자 한다. 후자는 실용번역의 영역에서 도출된 모델들이 문학번역의 평가에 적용가능인가에 대해 의문을 제기할 수도 있겠다. 물론 번역학 내부에서 평가 모델을 제시한 이론가들이 상당부분 실용번역의 평가를 그 출발점으로 하고 있음은 부정할 수 없다. 그러나 우리는 이러한 이론적 성찰들이 문학영역에서의 적용가능성을 배제하고 있지는 않음에 주목하고자 한다. 번역품질평가에 대한 이론적 담론이 척박한 국내의 상황에서는 기존의 담론들 속에서 유의미한 것들을 끌어내어 적용해 보고, 그 한계점을 보완해 가는 방식으로 작업하는 것만이 해답일 것이다.

2. 번역품질평가의 주요 경향

번역품질평가와 관련된 담론들을 분류하는 방식은 학자들마다 다르다(참고; Waddington 2001: 311, House 2001: 244-247, Williams 2001: 329-335, Colina 2008: 99-114). 본고에서는 번역학 내부에서 개진되고 있는 번역품질평가 담론들 중 몇 가지 주요한 경향들을 추려서 구체적 예와 함께 소개하고자 한다.

2.1. 주관적, 직관적 접근

번역을 개인의 주관적 해석이나 예술적, 문학적 해석적 직관을 토대로 한 창작물로 간주할 경우, 번역에 대한 평가 역시 종종 평가자의 경험이나 직관, 혹은 주관적 느낌을 토대로 이루어지는 것이 당연하다는 결론에 이르게 된다. 어떤 번역이 ‘원문의 맛을 살리는 번역’, ‘자연스러운 번역’, 혹은 ‘매끄럽지 못한 번역’이라고 평가할 때, 이것은 평가자가 ‘좋은 번역’에 대해 가지고 있는 인식을 토대로 이루어지는 주관적, 직관적 평가라고 볼 수 있다. 하우스는 평가자의 언어적, 문화적 직관이나 경험을 토대로 하고 있는 이러한 접근들을 ‘심리

주의적 관점(Mentalist view)'으로 분류하며, 번역물 평가와 관련된 논의에는 부적절함(inappropriate) 것으로 평가한다(House 2001: 244). 번역학이 엄밀한 '경험과학'으로 성장해야 한다고 믿는 몇몇 학자들도 이러한 담론들을 종종 자의적, 주관적, 처방적이라는 이유로 폄하한다.

그러나 번역에 대한 평가와 비평은 번역평가에 대한 이론적 논의가 활성화되기 훨씬 이전부터 무한히 다양한 방식으로 이루어져왔다. 예를 들어 번역교육기관에서 이루어지는 번역학습자들에 대한 번역능력평가는 명시적 기준 없이 수행되는 '자의적' 판단인 것처럼 보여도 현장에서 누적된 수년간의 경험과 노하우를 바탕으로 이루어지는 '전문가의 판단'이다. 우수 번역물의 선정 작업, 번역사의 교육 과정에서도 우리는 명시적 기준 없이도 무엇이 '좋은' 번역이고 누가 '훌륭한' 번역사인지를 판단해 왔다. 그 과정에서 축적된 노하우나 직관이 반드시 폐기되어야 하는 대상은 아니다. 문제는 이러한 직관과 노하우 자체에 있는 것이 아니라, 이것들이 충분히 종합, 점검, 체계화되지 못한 채 단순히 '누적'되어 있는 현 상황에 있다.

스타이너(Steiner 344)는 번역학 내부의 논의를 특징짓는 두 축으로 현장의 구체적 사례에서 출발하여 이론화작업에 이르는 본질주의적, 경험주의적 연구와 추상적 이론에서 출발하여 현장의 구체적 실무에 적용시키고자 하는 비본질주의적, 이론적 연구로 구분하며, 전자를 상향식(Bottom up) 접근으로, 후자를 하향식(Top down) 접근으로 규정한다. 이를 번역품질평가에 적용시킬 경우 이론의 장에서 마련된 개념적 틀이 실무현장에서 적용되는 방식(Top down)과, 실무적으로 평가가 어떻게 수행되어 왔는지를 관찰하는데서 출발하여 이를 모델화하는 방식(Bottom up)으로 구분할 수 있겠다. 예를 들어 실제 번역교사들이 번역물을 어떤 방식으로 평가하는지를 파악하기 위해 경험적 데이터를 수집하여 분석한 후, 그 결과를 바탕으로 번역평가모델을 제안하는 식의 접근은 상향식이라고 할 수 있겠다. 이러한 방식은 번역이론이 실제 번역사들이 현장에서 어떤 방식으로 번역하고 있는지를 관찰하는 것에서 출발해야 한다고 주장하는 모삽과 같은 학자들의 관점과도 일맥상통한다(Mossop 1998: 46). 특히 번역품질평가의 경우는 이론이 실무의 속도를 따라가지 못하는 면이 없지 않고, 따라서 번역품질모델 구상의 초기단계에서는 이러한 상향식 접근이 활용될 만하다.

2.2. 독자반응중심 접근(response-based approach)

번역물의 품질을 평가할 때 해당번역물이 독자에게 일으키는 반응이나 효과를 중심으로 평가하는 방식은 그 역사가 오랜 것으로, 좋은 번역이란 원문과 동일한 ‘효과’를 유발하는 번역이라는 전제로부터 출발한다. 성경번역의 경험을 이론화한 나이다(Nida 166)가 제안한 역동적 등가(dynamic equivalence)는 이러한 전제를 개념화한 것에 다름 아니다. 형식적 등가(formal equivalence)의 반대 개념으로서의 역동적 등가란 원문의 메시지가 번역된 방식이 번역텍스트의 수용자들로 하여금 원문독자와 유사한 반응(response)을 일으키는 방식으로 번역된 텍스트가 가지는 특징으로 정의된다. 이러한 접근이 품질평가에 적용될 경우, 이는 번역문에 대한 독자의 행태를 평가 기준으로 삼는다는 점에서 ‘행태주의적 접근(behavioristic approach)’으로도 정의된다(House 2001: 244). 이러한 접근에서는 대부분 ‘가독성’과 ‘정보성’을 갖춘 번역을 좋은 번역으로 간주하며, 따라서 나이다와 테이버(Nida & Taber 170)에서는 번역문의 가독성을 ‘빈칸완성테스트(cloze test)’를 통해서 측정해 볼 것을 제안하기도 한다. 예를 들어 번역 텍스트에서 다섯 단어마다 한 단어씩을 지우고 그 단어를 맞추도록 하여 맞추는 빈도가 높을수록 가독성이 높은 것으로 평가하는 식이다.

이러한 접근의 한계는 명백하다. 모든 번역이 정보성이나 가독성의 잣대로 평가되기는 어려우며, 설령 이 두 가지 기준만을 적용한다 하더라도, 정보성과 가독성을 정확히 측정해 내기는 어렵다. 그리고 독자의 반응(response)이라는 것 역시 몇몇 드문 경우를 제외하고는 ‘균질’하지 않다. 인터넷 서점 및 출판사의 인터넷 홈페이지를 통해 수집한 번역비평을 분석해 보면 동일한 번역에 대해서도 독자들의 반응이 무한히 다양한 방식으로 드러날 수 있음을 확인할 수 있다(이상원 2006). 또한 독자의 반응은 정보전달을 위주로 하는 특정유형의 텍스트, 혹은 특정 독자들(ex. 아동)을 대상으로 특정 목적(ex. 난해한 원전의 대중화)의 번역에는 적용될 수 있으나, 번역일반에 대해 유일한 기준으로 적용되기에는 무리가 있다. 그러나 독자반응중심 접근은 오랫동안 ‘원전에 대한 충실성’ 중심으로 전개되어 왔던 품질논의의 무게중심을 도착어 텍스트와 도착어 문화, 그리고 최종 수용자(receptor) 쪽으로 이동시켰다는 점에서 큰 의미가 있다. 실제로 번역텍스트의 최종독자들이 번역에 대해서 내리는 평가는 독자들이

번역으로부터 무엇을 기대하고 있으며 어떤 것을 좋은 번역으로 받아들이는지를 파악하는데 중요한 자료가 된다. 특히 번역의 목적이 원전에 대한 접근을 용이하게 만들어 대중화시키는데 있을 경우, 독자반응을 평가하여 이를 번역과정에 반영하기 위한 노력도 함께 이루어져야 할 것이다.

2.3. 언어학적 접근

이는 비교언어학, 대조언어학, 텍스트언어학, 코퍼스언어학, 비교문체론 등 주로 언어학의 영역에서 축적되어 온 이론적 연구들을 기반으로 번역품질 평가에 접근하는 일련의 시도들을 총칭한다. 이러한 작업들은 번역학의 탄생 초기부터 현재까지 꾸준히 이어져 왔다. 초기 품질평가 이론의 대표적인 예로 텍스트언어학을 번역품질평가에 접목시킨 하우스의 평가모델을 예로 들 수 있다. 번역품질평가에 대한 최초의 대대적 이론화 작업을 수행한 하우스의 평가모델은 원문텍스트의 자질들을 텍스트언어학적 개념들을 통해 분석하고 이를 바탕으로 원문이 어떤 방식으로 번역되어야 하는지(overt/covert)를 확인한 후, 해당 번역문이 실제로 어떻게 번역되었는지에 따라 번역의 품질을 평가하는 것을 주 내용으로 한다(House 1997). 이 밖에도 비교문체론 관점에서 번역품질평가를 시도한 다블네는 의미, 관용어, 어조의 존중, 문화적 요소, 인유(Allusion), 저자의 의도, 독자 등 7개 기준을 설정하고 이러한 기준들을 충족시키는지 여부로 번역의 품질을 판가름 할 수 있다고 설명한다(Darbelnet 1977). 그러나 이러한 평가모델들은 체계적이고 객관적인 평가의 기준과 방식을 제안했다는 점에서 의미는 있으나, 실무적으로 적용하려면 해당 모델에 대한 상당한 이론적, 개념적 이해를 요한다는 점에서 한계가 있다. 이는 하향식으로 제시된 평가모델에서 종종 드러나는 한계이다. 번역품질평가에 대한 이론적 연구들은 그 특성상 실제 평가 작업에서 활용되고 보완되지 못하면 크게 의미가 없다. 그런데 언어학적 접근들은 종종 번역모델의 적용가능성(applicability) 측면을 간과하였고, 이는 해당모델의 대중화를 어렵게 만든다.

언어학적 개념들을 활용하되, 실무적 적용가능성을 높인 평가시스템의 대표적인 예로 캐나다 정부산하 기관인 번역국(Translation Bureau)에서 번역품질의 체계적 관리를 위해 구상한 캐나다 언어품질평가시스템(SICAL, Système

canadien d'appréciation de qualité linguistique)을 들 수 있다. SICAL은 캐나다 번역국에서 매년 생산되는 엄청난 물량의 번역 품질을 효율적으로 관리하기 위해 고안된 것으로 여러 차례의 수정보완을 거쳐 최종적으로는 아래와 같은 단순한 오류구분 기준을 채택하게 되었다. 번역상의 오류를 크게 언어오류(language error)와 전달오류(transfer error)로 구분하고 이를 그 경중에 따라 사소한 오류와 심각한 오류로 다시 구분하여 오류의 개수에 따라 번역물의 품질 등급을 결정하는 것이 SICAL의 골자이다.

〈표1: SICAL의 번역품질평가 기준표, Larose 1998: 12〉

		400 단어당 오류 갯수	
		심각한 오류 (Major error)	사소한 오류 (Minor error)
A	우수	0	0-6
B	양호	0	7-12
C	감수 요	1	13-18
D	불량	1+	18+

위의 표에서 보는 바와 같이 번역의 최종 품질을 평가하는데 가장 결정적인 요소는 중요한 오류(Major error)의 수이다. 따라서 이 모델의 핵심은 중요한 오류와 사소한 오류를 각각 어떻게 정의할 것인가의 문제였다. 그런데 정작 SICAL이 지속적으로 사용되지 못한 이유는 보다 ‘정치적’ 차원에 있다. SICAL은 절대적 품질(absolute quality)이 아닌 상대적 품질(relative quality), 다시 말해 일정 수의 사소한 오류를 허용하는 평가 모델이다. 그런데 1995년 캐나다 번역국이 민간부문의 경쟁을 시작하면서 경쟁력 강화를 위해 상대적 품질 개념을 폐기하고 절대적 품질(error-free translation)을 채택하게 되었고 이로 인해 SICAL의 기본원칙 자체에 대해 문제가 제기되었던 것이다(Williams 2001: 330-331). 그러나 SICAL은 국가차원의 체계적 품질관리 모델로서 거의 유일하며, 실제 현장에서 활용되고, 그 과정에서 보완된 모델이라는 점에서 큰 의미가 있다.

기존의 직관적, 자의적 평가로부터 벗어나기 위해 언어학과의 협력 하에 시도된 오류의 범주화, 계량화 작업은 예상했던 것만큼 큰 효과를 거두지는 못하

였다. 번역의 품질평가라는 것이 전적으로 객관화될 수 없는 작업일 뿐 아니라, 아무리 이론적으로 완성도 있는 평가모델이라 해도 실제 운용의 문제가 충분히 고려되지 않을 경우, 그것은 이론적 사변으로만 남게 됨을 다시 한 번 확인할 수 있다. 더욱이 단어나 문장 단위에서의 오류 분석과 분류에 치중하다 보니 번역텍스트의 담화적 차원, 기능적 차원이 종종 무시되었다는 점은 언어학적 접근들이 끝내 극복하지 못한 한계라고 볼 수 있다.

2.4. 기능주의적 접근(functionalist approach)

독일을 중심으로 전개된 기능주의적 접근에서는 번역의 ‘기능’을 번역품질평가의 핵심어로 선택한다. 번역의 ‘충실성’이나 ‘가독성’은 더 이상 논의대상이 아니며, 번역문의 좋고 나쁨은 번역문이 주어진 기능을 얼마나 훌륭히 수행해 내는가에 따라 결정된다. 1970년대 말에서 80년대 초 독일을 중심으로 형성된 기능주의 학파는 이후 80년대에 이르러 기술번역학(Descriptive Translation Studies)과 함께 도착어 중심적 번역담론(Target-oriented TS)을 주도하게 된다. 이들의 주장은 스키포스 이론(Skopos theory)으로 정리되는데, 번역이 도착어 맥락에서 어떤 목적(Skopos)을 가지는가에 따라 번역의 방식, 전략을 비롯한 모든 것을 규정한다는 것이 그 핵심이다. 이들에게 원문은 더 이상 ‘성스러운’ 것이 아니며, 목적에 따라 자유로이 조정, 가감할 수 있는 대상에 지나지 않는다. 따라서 이들의 원문에의 충실성보다는 도착어의 완결성, 언어적 표현력, 전문용어의 적절한 사용 등을 평가의 중요한 기준으로 삼게 된다. 스키포스 이론은 번역이 특정상황에서 구체적 목표와 수요를 충족시키기 위해 이루어진다는 너무도 당연한, 그러나 종종 간과되었던 사실을 상기시켜주었다는 점에서 큰 의미가 있다. 그러나 번역품질평가에 적용할 경우 번역의 목적이 제대로 충족되었는지 여부를 구체적으로 어떻게 평가할 것인가의 문제, 혹은 문학텍스트의 번역과 같이 특정한 ‘기능’이나 ‘목적’을 가지지 않는 것으로 보이는 작업에의 적용가능성 등이 과제로 남아 있다.

이제까지 번역품질평가와 관련하여 번역학 내부에서 이루어진 연구들을 총 네 가지로 구분하여 살펴보았다. 번역학 내부에서 번역품질평가가 본격적으로

논의되기 시작한 것은 두 개의 번역학 전문학술지의 공이 크다. *The translator*지는 2000년도에 발행된 6권 2호의 제목을 평가와 번역(Evaluation and Translation)으로 결정하고 전체 지면을 모두 번역물의 평가 문제에 할애하였다. 마이어(Maier 137)가 서론에서 밝힌 바에 따르면 이 특별호의 목적은 ‘평가에 대한 해결책이나 정의를 제시하기보다는 현재의 평가 방식의 문제점을 다양한 관점에서 다른 연구들을 취합, 검토하고 이를 통해 새로운 출발을 시도’하는 것이었다. 실제로도 이 특별호에서 다루어진 주제 및 관점의 다양함을 일별해 보면, ‘해결책’이나 ‘정의’는 여전히 요원한 상태임을 확인할 수 있다. 이러한 느낌은 다음해인 2001년, 또 다른 번역학 전문 학술지인 *Meta*에서 기획한 번역평가 관련 특별호를 일별한 후에도 크게 달라지지 않는다(참고 *Meta* 46 (2)). 여기서도 통역에서의 품질평가가 추가된 것이 차이일 뿐, 평가에 대한 다양한 실험적 접근들이 개괄적으로 제시되어 있는 수준에 그치고 있기 때문이다.

3. 기존 번역품질 평가 담론의 문제점 혹은 간과된 요소들

번역학 내부에서 개진된 번역품질평가 관련 논의는 무엇이 좋은 번역인가에 대한 막연하고 직관적인 논의들을 이론화, 공론화시키는데 기여하였다. 그러나 현재까지의 품질평가 관련 논의들이 실무 현장과 얼마나 효율적으로 연계되어 왔으며, 현장에서의 번역품질평가 관행을 얼마나 많이 변화시켜왔는지에 대해 묻는다면 그 대답은 부정적일 수밖에 없다. 예를 들어 얼마나 많은 번역교육 기관에서 하우스의 품질평가 모델(Housian model)을 실제로 사용하고 있으며, 얼마나 많은 국가에서 번역품질관리를 위하여 SICAL을 참고하였는가? 번역물 품질평가에 대한 개념적, 이론적 논의들은 평가의 실제(practice)와 상당한 거리를 두고 진행되어 온 것이 사실이며, 따라서 번역학계 내부에서 제시된 평가모델들이 실무 현장에서 충분히 적용, 검증, 보완되었다고 보기는 어렵다. 이는 기존의 이론적 논의들이 평가의 실제에서 발생하는 실무적 제약과 문제들을 충분히 고려하지 않은 데에 기인한다. 따라서 향후 번역품질 담론에서 추가적으로 반드시 논의되어야 할 몇 가지 사항들을 정리해 보면 아래와 같다.

3.1. 평가의 궁극적 목적의 문제

기존의 번역품질평가 모델들은 대체로 평가의 기준, 오류의 종류, 비중 등을 구분하는데 치중한 반면, 평가의 목적에 따라 평가 방식이 달라진다는 점을 간과한 면이 있다. 예를 들어 동일한 고전문학작품을 평가한다 하더라도, 평가의 목적에 따라 평가의 기준이나 오류의 비중이 달라질 수 있다. 해당 평가가 번역자의 번역능력을 제고하거나 품질을 개선하는 등의 교육적 목적을 가질 경우, 평가자는 일종의 교육자(educator)의 역할을 수행하게 되며 따라서 형성평가(formative assessment)를 수행하게 된다. 이러한 경우 평가기준이 사전에 마련되어 평가자와 번역자가 이를 공유하여야 하며, 평가자가 바람직한 것으로 여기는 특정 방식으로 번역이 이루어지도록 사전, 사후적인 교육이 필요할 뿐 아니라, 평가 결과를 번역자와 공유하기 위한 적절한 피드백 메커니즘도 마련되어야 한다. 특히 좋은 번역을 구성하는 다양한 항목들을 평가할 수 있는 항목별 평가(componential evaluation)가 이루어져야 세부적이고 설명력 있는 피드백(constructive feedback)이 제공될 수 있다.

반면, 공신력 있는 국가기관이나 혹은 해당분야의 전문성을 가진 평가자가 일종의 ‘심판’으로서 번역물에 대한 최종적 ‘판결’이나 당락을 결정할 경우, 이는 총괄평가(summative)적 성격을 띠게 되고 따라서 교육적 평가에 비해 비교적 단순한 방식의 총체적 접근(holistic approach)이 선택될 수도 있다.

따라서 ‘무엇을 위한 평가인가’가 충분히 논의되지 않은 상태에서 평가의 기준이나 방식에 대해 논하는 것은 무의미하다.

3.2. 적용가능성(applicability)의 문제

앞서 언급한 대로 번역이론가들이 주축이 되어 평가모델을 고안할 경우, 종종 간과되는 측면 중의 하나가 바로 평가모델의 실제 적용가능성(applicability)의 문제이다.

구아텍(Gouadec 1981)이 캐나다 번역국으로부터 연구를 위탁 받아 제안한 번역물실증평가체계(SEPT, *Système d'évaluation positive des Traductions*)가 그 대표적 예로 라로즈는 이를 두고 ‘평가모델의 이론적 가치와 그 적용가능성은 반비례한다’고 지적할 정도였다(Larose 1998: 12). 구아텍이 제시한 평가모델은

크게 세 단계로 구성되는데, 첫째, 각각의 오류의 성격을 분석하고 둘째, 각각의 오류가 텍스트 전체에 어느 정도의 영향을 주는지에 따라 오류의 비중을 계산한 후, 셋째, 다시 그 결과를 텍스트의 난이도를 감안하여 보정하는 식이다. 문제는 구아텍이 코드화한 오류의 종류가 총 675종(어휘적 오류 300종 + 통사적 오류 375종)에 이른다는 것이다. 평가모델이 이렇듯 고도로 복잡할 경우, 평가자들의 교육 문제, 이에 소요되는 시간과 예산의 문제가 제기되고 궁극적으로는 이러한 방식의 평가가 실무적으로 적용가능한가의 문제가 제기될 수밖에 없다. 구아텍의 평가모델이 보급되지 못하고 사장된 것은 충분히 예측가능한 일이었다.

평가의 효율성을 위하여 종종 번역물의 전체가 아닌 일정 부분을 선택, 평가하는 샘플링 평가 방식이 적용되기도 한다. 이 역시 평가모델의 효율성을 높이기 위해 고안된 방안의 하나로, 앞서 언급한 SICAL 역시 이러한 샘플링 방식을 채택하였다. 그러나 샘플의 크기, 선정방식, 대표성 문제 등 또 다른 문제들이 제기되는 것을 피할 수 없다.

3.3. 원문의 위상 문제

번역품질평가를 논할 때, 실용번역과 문학번역을 구분하거나 혹은 텍스트 유형(text type)에 따라 번역물의 평가기준이 달라진다고 보는 것이 일반적이다. 예를 들어 라이스(Reiss 2000)는 텍스트 유형을 크게 형식위주(Form-focused), 내용위주(Content-focused), 호소위주(Appeal-focused) 텍스트로 분류한다. 시나 문학 작품 등 형식위주 텍스트의 경우는 언어의 형식, 즉 미학적 예술적 측면이 중요하고 따라서 번역문의 표현 측면이 가장 중요하게 평가되어야 하고, 정보 전달을 중요시하는 내용위주 텍스트는 원문의 내용이 얼마나 잘 전달되었는지가 평가의 기준이 되어야 한다는 것이다. 반면 호소 위주 텍스트의 경우는 특정 정서를 전달하는 것이 중요하며, 독자에게 동일한 효과(effect)를 유발하는 것이 중요하다. 라이스의 관점에서는 나이다의 역동적 등가 개념은 호소위주 텍스트에만 적용가능하다. 그러나 실제로 텍스트의 유형을 분류하는 것은 쉽지 않으며, 하나의 텍스트가 동시에 여러 측면을 복합적으로 가지고 있는 경우가 많음을 감안할 때, 텍스트 유형에 따른 분류는 그 이론적 기여에도 불구하고 실무적

으로는 상당히 한계가 있어 보인다.

오히려 우리는 번역품질평가의 방식을 규정하는 중요한 요소 중 하나로 ‘원문의 위상’을 생각해 볼 수 있겠다. 번역의 목적(skopos)에 의해 번역방식이 규정되고, 원문은 단지 정보의 제공일 뿐 엄격히 존중해야 할 대상이 아니라고 주장하는 스크포스 이론가들의 논리가 문학번역 진영으로부터 비난받는 이유는 원문텍스트의 위상이나 비중이 모든 종류의 번역에서 동일하지 않다는 사실을 이들이 간과했기 때문이다. 정보전달을 위주로 하는 실용텍스트의 번역의 경우 원문에 대한 숭배(cult)는 존재하지 않는다. 반면 고전작품이나 종교경전들의 경우, 원문은 번역과정에서 보존하고 재구성해내야 할 다양한 차원들을 담고 있는 일종의 ‘신성한 텍스트(sacred text)’들이며, 따라서 이러한 원전의 번역에 대한 평가 작업은 어쩔 수 없이 문헌학적, 고증적, 주해적 작업을 수반한다. 그런데 바로 이러한 이유 때문에, 소위 신성한 텍스트의 평가 과정을 체계화하기가 어려워진다. 일대일 대응관계로 번역 가능한 전적으로 기술적인 텍스트들과는 달리, 이러한 텍스트들은 여러 가지의 해석 가능성이 열려 있는 텍스트들이기 때문이다. 따라서 여기서는 엄밀한 의미에서의 ‘오류’나 ‘오역’의 개수 문제를 경직된 방식으로 논하는 방식으로 평가가 이루어지기는 어려우며, 다양한 ‘해석의 차이’, ‘관점의 상이함’에 대한 고찰이 평가 과정에서 함께 다루어져야 한다.

3.4. 평가자의 문제

오랫동안 번역에 대한 논의가 번역행위의 결과물로서의 번역에만 관심을 두고, 정작 번역을 생산하는 주체인 번역사들을 간과해 온 것과 마찬가지로, 번역 품질평가에 관한 논의에서도 종종 평가의 핵심인 ‘평가자’들이 간과되곤 한다. 예를 들어 서로 다른 평가자들에 의해 이루어진 평가 간에 어느 정도의 편차가 존재하는지를 보여주는 평가자간 신뢰도(inter-rater reliability)는 평가모델 자체의 신뢰도와 직결된 요소이다. 평가자간 신뢰도를 높이기 위해서 평가자는 해당분야에 대한 전문지식 뿐 아니라, 평가 모델에 대한 충분한 이해를 가지고 있어야 한다. 평가모델이 아무리 정교해도, 평가는 결국 평가자에 의해 수행되는 것이며, 평가자 그룹을 어떻게 구성하느냐에 따라 평가의 결과도 충분히 달라질 수 있다.

평가 작업의 성격에 따라 이상적 평가자의 자격조건에 대한 심도 깊은 논의도 필요하다. 예를 들어 번역능력과 평가능력을 동일한 것으로 보고, 번역가에게 요구하는 것과 동일한 능력을 평가자에게 요구할 것인가? 아니면 번역물의 최종 독자와 유사한 조건을 가진 평가자를 선택할 것인가? 혹은 평가 대상이 되는 번역물의 특수성을 감안한 추가적 역량을 요구할 것인가?

또한 평가능력을 향상시키기 위해 활용할 수 있는 것은 무엇인지 고민해 볼 필요도 있다. 예를 들어 특정 장르의 텍스트를 평가할 경우, 평가자들이 평가에 활용 가능한 문서들을 코퍼스나 DB로 구축하여 활용할 수 있도록 하는 방식도 검토 가능하다.¹⁾

3.5. 과정(process) 속에서의, 과정으로서의 평가

보통 번역품질평가는 번역물이 생산된 이후, 사후적으로 개입하여 해당번역물의 품질을 평가하는 작업으로 이해된다. 그러나 번역품질평가는 번역품질보증(quality assurance)이라는 상위의 프로세스를 구성하는 한 단계로 인식될 때만이 효율적으로 기능할 수 있다. 여기서 번역품질보증이란 번역의 결과물만이 아닌 해당번역물의 생산에서 최종인도에 이르기까지의 전 과정을 관리하는 것을 말한다(Mossop 2001: 93). 즉, 번역자의 선정, 번역자에 대한 대우, 번역기한의 설정 등의 단계에서부터 출판(혹은 납품)단계에 이른 후 사후 모니터링에 이르기까지 최적의 번역물을 생산하는데 필요한 전반적 조건에 대한 성찰 속에서만이 효율적인 품질평가방식이 구상, 실현될 수 있는 것이다. 번역품질의 문제를 ‘시스템’ 개념으로 접근해야 하는 이유도 여기에 있다.

또한 최종 번역물의 평가 작업 자체도 개별적, 일회적 평가행위로서가 아닌 ‘과정’으로 인식하는 것이 중요하고 따라서 ‘평가 작업의 흐름도’가 마련되어야 한다. 예를 들어 나이다는 대규모의 인력이 투입되는 공동번역의 경우 집필위원회(editorial committee), 검토위원회(review committee), 자문위원회(consultative committee) 등의 단계로 작업이 이루어진다고 설명한다(Nida 247-251). 집필위

1) Bowker(2000)에서는 학생들의 번역품질평가에 코퍼스를 활용한 사례를 소개하고 있다. 또한 Bowker(2001)에서는 번역교사들이 번역교육과정에서 코퍼스를 활용하는 방안을 제안하고 있다.

원회가 번역의 초안을 작성하고 검토위원회에서는 해당분야의 전문적 역량을 가진 인력들이 번역의 전반적 내용을 검토, 필요한 제안을 하며, 마지막으로 자문위원회는 번역의 생산과정에는 참여하지 않으면서 해당번역의 수용여부를 최종적으로 판단하는 역할을 수행한다. 이렇듯 평가 작업을 개별 평가자에 의한 일회적 작업이 아닌 하나의 흐름도로 이해하고 평가를 포함한 번역작업 전체를 염두에 둔 번역품질평가 모델이 필요하다.

4. 나오는 말

이상에서 번역품질평가와 관련된 주요 접근들을 개괄하고, 향후의 번역품질 평가 논의에서 감안되어야 할 몇 가지 요소들을 살펴보았다. 현재로서는 번역의 품질에 관한 이론적 담론들이 실무를 통해 보완되고, 현장에서 축적된 노하우가 효율적으로 이론화되는 단계에 이르렀다고 보기는 어려운 것이 사실이다. 우리는 번역품질에 대해 개별 사례위주로 단발적인 논의를 축적한다고 해서 그것이 번역품질에 대한 체계적 담론으로 이어지는 않음을 경험을 통하여 깨달았다. 우리의 번역현실을 개탄하고 열악한 번역 여건을 비판하는 것만으로는 번역품질의 향상이 이루어지기 힘들다는 것은 자명한 사실이다. 그러나 추상적이고 막연한 상태에 머물러 있던 번역품질관련 논의들이 비로소 구체적 시도들로 이어지기 시작하고 있는 것은 반가운 사실이다. 수많은 비판에도 불구하고 영미언에서 수행한 번역평가 사업이 의미 있다고 판단되는 이유는 이러한 작업을 통해 평가 작업이 내포하고 있는 복합적이고도 모순적인 다양한 측면들이 드러나고, 그에 대한 구체적이고 진지한 성찰들이 촉발되었기 때문이다.

다만, 명저의 번역이나 고전번역 등을 통해 주도되고 있는 국내의 품질 논의들이 번역학 내부에서 미미하게나마 진행되고 있는 실무적 논의들과 보다 효율적으로 연계된다면 논의의 활성화에 도움이 될 것이다. 또한 고전작품을 포함한 문학작품을 중심으로 한 평가 논의를 넘어서서 의료서비스의 통역이나 번역, 법정 통역 등 국내에서 급격히 발전되고 있는 지역사회통번역(*community interpreting*)에까지 번역품질에 대한 성찰이 확장되기를 기대한다. 주로 사회적 약자들을 대상으로 이루어지고 있는 지역사회통번역의 영역에서 통역 및 번역

의 열악한 품질 문제는 학문적 문제를 넘어 사회적 불평등의 문제, 더 나아가 인권의 문제와도 연결되기 때문이다(참고 Colina 2008). 이런 측면에서 볼 때, 번역 품질평가라는 영역은 번역학이 이론과 실천, 경험과 반성의 화해를 가장 훌륭히 실험할 수 있는 장이라고 판단된다.

참고문헌

- 영미문학연구회 번역평가사업단 (2005) 『영미명작, 좋은 번역을 찾아서』, 서울: 창비.
- 영미문학연구회 번역평가사업단 (2007) 『영미명작, 좋은 번역을 찾아서 II』, 서울: 창비.
- 이상원 (2006) 『한국출판번역독자들의 번역평가규범 연구』, 파주: 한국학술정보(주).
- Baker, Mona (1996) 'Linguistic and cultural studies: Complementary or competing paradigms in Translation Studies?'. Angelika Lauer, Heidrun Gerzymisch-Arbogast, Johann Haller and Erich Steiner (Eds). *bersetzungswissenschaft im Umbruch: Festschrift für Wolfram Wilss zum 70 Geburtstag*. Tübingen: Gunter Narr (pp. 9-19).
- Bowker, Lynne (2000) 'A Corpus-based approach to evaluating student translations', *The translator* 6(2): 183-210.
- Bowker, Lynne (2001) 'Towards a methodology for a corpus-based approach to translation evaluation', *Meta* 46(2): 345-363.
- Chesterman, Andrew (1998) 'Causes, Translations, Effects', *Target* 10(2): 201-230.
- Colina, Sonia (2008) 'Translation Quality Evaluation: Empirical Evidence for a Functionalist Approach', *The Translator* 14(1): 97-134.
- Darbelnet, Jean (1977) 'Niveaux de traduction', *Babel* 23(1): 6-17.
- Gouadec, Daniel (1981) 'Paramètres de l'évaluation des traductions', *Meta* 26(2): 99-117.

- House, Juliane (1997) *Translation quality assessment : A model revisited*,
Tubingen: Gunter Narr Verlag.
- House, Juliane (2001) 'Translation quality assessment: Linguistic description
versus social evaluation' *Meta* 46(2): 297-301.
- Larose, Robert (1998) 'Méthodologie de l'évaluation des traductions', *Meta*
41(2): 1-24.
- Maier, Carol (2000) *Evaluation and Translation. Special Issue. The Translator:
Studies in Intercultural Communication* 6(2). St. Jerome.
- Mossop, Brian (1998) 'The workplace procedures of professional translators'.
In A. Chesterman et al (Eds.) *Translation in context* (pp. 39-48).
- Mossop, Brian (2001) *Revising and editing for translators*. Manchester &
Northampton: St. Jerome.
- Nida, Eugene A (1964) *Towards a science of Translating: With special
reference to principles and procedures involved in bible translating*,
Leiden: E. J. Brill.
- Nida, Eugene A, & Taber, Charles R (1982) *The theory and practice of
translation*, Leiden: E. J. Brill.
- Reiss, Katharina ([1971]2000) *Translation criticism: The potentials &
limitations: Categories and criteria for translation quality assessment*.
Rhodes, E.(Trans.). *Möglichkeiten und Grenzen der Übersetzungskritik*,
Manchester: St. Jerome.
- Steiner, Erich (2001). 'How (translated and otherwise interlingual) texts works
is our way into what, why and to what effects', *Target* 13(2): 343-348.
- Waddington, Christopher (2001) 'Different Methods of evaluating student
translations: The question of validity' *Meta* 46(2): 311-325.
- Williams, Malcolm (2001) 'The application of argumentation theory to
translation quality assessment', *Meta* 46(2): 326-344.

[Abstract]

Translation Quality Assessment: review and suggestions

Lee, Hyang
(Korea University)

Recently, there has been growing interest in the quality assessment of translation conducted in Korea, and serious discussions regarding translation quality, focused mainly on literary works, are being carried out. Within this context, this article outlines and examines some major trends of the theoretical approaches of translation quality assessment that have been discussed, and then describes some elements that should be taken into account in future discussions of translation quality. The first section of the article provides a brief introduction to the subjective/intuitive, response-based, linguistic, and functionalist approaches of translation quality assessment, and explains the characteristics and limitations of each approach. The following section presents five aspects that have been overlooked in the existing evaluation models: the purpose of the evaluation, the applicability of evaluation models, the nature of source text, the evaluator, and the entire process of quality assurance.

▶ Key Words: Translation quality assessment, subjective/intuitive approach, response-based approach, linguistic approach, functionalist approach.

이 향

고려대학교 불어불문학과 연구교수

leehyangmarina@gmail.com

관심 분야: 번역 품질 평가, 연구방법론, 번역 철학

논문투고일: 2010년 4월 29일

심사완료일: 2010년 6월 09일

게재확정일: 2010년 6월 10일