Discourse marker *well* in the spoken English of Taiwanese learners

Lan-fen Huang

Assistant Professor, Language Centre
Shih Chien University, Kaohsiung, Taiwan

2 September 2015

Outline

1. ‘Well’ in spoken English
2. Brief introduction to LINDSEI
3. Data and methods
4. *Well* in *Taiwanese* Sub-corpus of LINDSEI
    4. 1 frequency
    4.2 position
    4.3 function
5. Discussion
6. Implications

Well in spoken English

‘Well’

as a marker of

FLUENCY or **DIS-FLUENCY**?

- Learners' L1s and their preference are attributing factors in the use of discourse markers (Gilquin & Granger 2015).
- 'Well' over-represented in LINDSEI-FR (Gilquin 2008) and LINDSEI-SW (Aijmer 2011) compared to the native counterparts in LOCNEC (De Cock 2004)
- **How about 'well' in LINDSEI-TW?**

over-representation or under-representation of 'well' in learners' speech--

Does it matter?

Brief introduction to LINDSEI

- **The Louvain International Database of Spoken English Interlanguage** (**LINDSEI**) version 1 (Gilquin et al. 2010), began in 1995, published in 2010 by the Centre for English Corpus Linguistics, Catholic University of Louvain, Belgium
- **11 sub-corpora of different L1s:** Bulgarian, Chinese, Dutch, French, German, Greek, Italian, Japanese, Polish, Spanish and Swedish (9 more, see Gilquin 2015)
- 544 informal interviews;

  An average of 1,949 tokens;

  Roughly 1-million tokens, 33% from interviewers, 66% from learners

Data & tools in this study

Learner data: LINDSEI-TW (Huang 2014)

Native speaker data: LOCNEC (De Cock 2004)

Tools: *WordSmith 6* (Scott 2012)

LINDSEI-TW

Corpus design criteria

- Speakers: 50 third- or fourth-year English majors
- Duration: 15 mins
- Interview contents (experience, country, film/play)
- Learner profiles
- Orthographically transcribed and marked up according to the transcription guidelines

Table 1. The Structures of the Taiwanese Sub-corpus of LINDSEI

Corpus		LINDSEI-TW
Recording dates		From 19 Nov 2012 to 3 Jun 2013
Composition of corpus	No. of interviews	50
	No. of tokens (Turns A & B)	110,280
	No. of tokens (Turns B only)	69,577
	No. of tokens per task (Turns A & B)	Set topics: 36,905 (33%) Free discussion: 60,307 (55%) Picture description: 13,068 (12%)
	No. of tokens per task (Turns B only)	Set topics: 25,969 (37%) Free discussion: 35,450 (51%) Picture description: 8,158 (12%)
	Total duration	12 hours 54 minutes
Interview	Average length (Turns A & B)	2,206
	Average length (Turns B only)	1,392
	Average duration	15 minutes 6 seconds
	Set topic	Country: 44% Experience: 34% Film/play: 22%
Learner	Average age	21.7
	Gender (percentage of female)	86%
	Average no. of years of English at school	9.38
	Average no. of years of English at university	3.22
	Average no. of months in English-speaking countries	2.81
	English proficiency (in CEFR levels) based on self-reported test results	B1: 13 (26%) B2: 17 (34%) C1: 19 (38%) C2: 1 (2%)
Interviewer	Gender (percentage of female)	16%
	Mother tongue (percentage of English NS)	70%

Adapted from Huang (2014, pp. 42-43)

Aim of the research

- To investigate the use of *well* in LINDSEI-TW and compare it with LINDSEI-FR (Gilquin 2008), LINDSEI-SW and LOCNEC (Aijmer 2011)

1. Frequency
2. Position
3. Function

Frequencies of *well* across LOCNEC, LINDSEI-FR, LINDSEI-SW and LINDSEI-TW

	LOCNEC	**LINDSEI-FR** (Gilquin 2008)	**LINDSEI-SW** (Aijmer 2011)	**LINDSEI-TW** (Huang 2014)
No. of interviews	50	50	50	50
No. of words (both turns of interviewers and interviewees)	170,533	149,127	102,131	110,280
No. of words (turns of interviewees)	125,666	94,406	71,853	69,577
Raw frequency of discourse marker *well* in turns of interviewees	529	1016	391	106
Relative frequency of discourse marker *well* (per 100,000 words)	415.25 (in Gilquin 2008) 420.96 (in Aijmer 2011)	1076.20	544.17	152.35
Log-likelihood score		+326.94	+14.59	-112.17
Significance*		$p < 0.0001$	$p < 0.001$	$p < 0.0001$

* Log-likelihood calculator, created by Rayson at http://ucrel.lancs.ac.uk/llwizard.html is used.

The distribution of *well* in learners in LINDSEI-TW

	File	Frequency	per 100,000	CEFR	Sub-total	%	Months Stayed in countries where English is spoken
1	TW001	14	20.12	C2	14	13	12.75
2	TW002	24	34.49	C1	42	40	12
3	TW004	4	5.75	C1			0
4	TW005	4	5.75	C1			84
5	TW010	1	1.44	C1			0.5
6	TW016	9	12.94	C1			1
7	TW021	1	1.44	B2	16	15	4.4
8	TW025	1	1.44	B2			0
9	TW030	1	1.44	B2			0
10	TW035	8	11.50	B2			0
11	TW037	5	7.19	B2			3
12	TW039	1	1.44	B1	34	32	0
13	TW040	33	47.43	B1			0
well (Overall)		**106**	**152.35**		**106**	**100**	

The distribution of the positions of *well* in LINDSEI-TW and LOCNEC

	LINDSEI-TW		LOCNEC[1]		X^2	Significance[2]
	n.	%	n.	%		
Utterance-initial	71	**67.0**	303	**57.3**	3.04532	n.s.
Utterance-medial	35	**33.0**	222	**42.0**	2.57449	n.s.
Utterance-final or alone	0	0	4	0.7	0.05089	n.s.
Total	106	100	529	100		

[1] The figures of LOCNEC are from Aijmer (2011, p. 235).
[2] The two-samples of SIGIL (SIGIL: Corpus Frequency Test Wizard) at http://sigil.collocations.de/wizard.html is used.

The distribution of the functions of *well* in LINDSEI-TW and LOCNEC

	LINDSEI-TW		LOCNEC [1]		X^2	Significance [2]
	n.	%	n.	%		
Speech management						
1. Choice	17	16.0	34	6.4	9.77917	p < 0.01
2. Change	20	18.9	138	26.1	2.09108	n.s.
3. Prospective (introducing a new turn)	43	40.1	107	20.2	19.13579	p < 0.001
4. Marking stages in a narrative	14	13.2	16	3.0	18.14328	p < 0.001
5. Quotative	1	1.0	49	9.3	7.31755	p < 0.01
Subtotal	**95**	**89.2**	**344**	**65.0**	**23.89200**	**p < 0.001**
Attitudinal						
6. Opinion	6	5.7	65	12.3	3.26625	n.s.
7. Disagreement	5	4.7	120	22.7	16.91255	p < 0.001
Subtotal	**11**	**10.4**	**185**	**35.0**	**23.89200**	**p < 0.001**
Total	106	100	529	100		

[1] The figures of LOCNEC are from Aijmer (2011, p. 248).
[2] The two-samples of SIGIL (SIGIL: Corpus Frequency Test Wizard) at http://sigil.collocations.de/wizard.html is used.

- **The prospective function**

  often occurs utterance-initially to give answers to the interviewer's questions

<A> (erm) . what kind of challenges do you think people have to face . when they're starting their lives in a new country </A>

*<B> (erm) . **well** if they're in . maybe America I'm guessing . (eh) they would have to face . (eh) . racism </B>*

(LINDSEI-TW TW005)

- ***Well* marking stages in a narrative**

overused by Swedish learners (Aijmer 2011) as well as German learners (Müller 2004). In contrast, it is underused by Taiwanese learners. 14 instances are found and 10 of them are from the same speaker, TW040, who gives 33 instances of *well* in the interview. The two instances of *well* highlighted in bold in the example below mark stages in the story of pictures.

<B> (em) cos she think well that's not me <laughs> that's not me </B>

<A> (mm) </A>

*<B> **well**: so: again (em) the man fix it . fix some s= fix many .. many: many details <laughs> </B>*

<A> (mhm) like what for instance </A>

*<B> like her hairs . her hair well and her . nose and her eyes everything is different not her: <laughs> and: . <X> and in the end **well** the woman show her friends </B>*

(LINDSEI-TW TW040)

- ***Well* mitigating speaker's disagreement**

  Well can be seen as a signal for correction of the preceding idea. In the example below, *well* is followed by a disagreement with the interviewer.

<A> <laughs> okay . yeah (em) .. so . yeah I I I noticed that . students from . your department from this school (em) . pretty much have quite (eh) good English abilities . and . in that case . in the future after you guys graduate (em) . what kind of jobs will you will you be able to find . if you have this kind of English abilities </A>

<B> ***well****: . actually I think that a lot of people speak good English <overlap /> these days </B>*

<A> <overlap /> yeah </A>

(LINDSEI-TW TW016)

Summary

1. The low incidence and the small number of learners who use *well* show that the learners in LINDSEI-TW underuse *well*, compared with native speakers in LOCNEC.
2. Both the NSs in LOCNEC and the Taiwanese learners use *well* in utterance-initial place in most cases (67% and 57.3% respectively)
3. In terms of functions, there are statistically significant differences in most functions, except the functions of change and opinion.

Implications

- The greater use of the functions for speech management may result from speakers' using English as a foreign language and the nature of spontaneous interviewing. This use does not generally create negative effects;

  but the underrepresentation of the attitudinal functions may suggest that Taiwanese learners sound unfriendly and hinder easy communication.

- Different mother tongues among learners require different pedagogical treatment. For Taiwanese learners some pedagogical intervention is suggested.

Implications

BUT

How to teach *well*?
What materials can be used?

Recommendations for teaching *well*

- **Starting point:** Grammar references, such as *Longman Grammar of Spoken and Written English* (Biber et al. 1999) and *Cambridge Grammar of English* (Carter & McCarthy 2006)
- **Consciousness-raising activities:** to train learners' ability to notice. Some published materials are available; e.g. *Touchstone 2* (McCarthy, McCarten, & Sandiford 2005)
- **Self-prepared handout:** Data in LOCNEC and LINDSEI-TW may also be used for teaching. For instance, *well* with negative response *no* may soften speaker's opinoins.

Example of *well* in LOCNEC

<B> I mean .. many people have said this that you know you wanna become an actor why don't you do the theatre studies <\B>

<A> mhm <\A>

<B> but it is a critical course <\B>

<A> yes <\A>

<B> and although [I I <\B>

<A> [you're not really interested in that <\A>

<B> **well n= no** it's not that I'm not interested . it's just .. well yes I suppose it is really . I'm not very good at it and I'm more interested in the practical side if it if it if it [was <\B>

<A> [being on stage <\A>

<B> if it was a course that looked at the script for .. looked at a[ei] script for a a certain amount of time and then performed it <\B>

(LOCNEC E04)

References

Aijmer, Karin (2011). *Well I'm not sure I think...The use of well by non-native speakers. International Journal of Corpus Linguistics, 16(2), 231-254.*

Biber, Douglas, Finegan, Edward, Johansson, Stig, Conrad, Susan, & Leech, Geoffrey. (1999). *Longman Grammar of Spoken and Written English. Essex: Pearson Education Limited.*

Carter, Ronald, & McCarthy, Michael. (2006). *Cambridge Grammar of English. Cambridge: Cambridge University Press.*

De Cock, Sylvie. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures, New Series 2, 225-246.*

Gilquin, Gaëtanelle (2008). Hesitation markers among EFL learners: Pragmatic deficiency or difference? In J. Romero-Trillo (Ed.), *Pragmatics and Corpus Linguistics: A Mutualistic Entente (pp. 119-149). Berlin, Heidelberg & New York: Mouton de Gruyter.*

Gilquin, Gaëtanelle, De Cock, Sylvie, & Granger, Sylviane (Eds.). (2010). *LINDSEI Louvain International Database of Spoken English Interlanguage. Handbook and CD-ROM. Louvain-la-Neuve: Presses universitaires de Louvain.*

Gilquin, Gaëtanelle, & Granger, Sylviane. (2015). Learner language. In D. Biber & R. Reppen (Eds.), *Cambridge Handbook of Corpus Linguistics*. Cambridge: Cambridge University Press.

Gilquin, Gaëtanelle. (2015, 1 April). LINDSEI Partners. Retrieved 19 April, 2015, from http://www.uclouvain.be/en-307845.html

Huang, Lan-fen. (2014). Constructing the Taiwanese component of the Louvain International Database of Spoken English Interlanguage (LINDSEI). *Taiwan Journal of TESOL, 11(1), 31-74.*

McCarthy, Michael, McCarten, Jeanne, & Sandiford, Helen. (2005). *Touchstone Teacher's Edition 2. Cambridge: Cambridge University Press.*

Scott, Mike. (2012). WordSmith Tools (Version 6). Liverpool: Lexical Analysis Software.

Acknowledgments

My thanks go to:

- National Science Council, Taiwan (grant number NSC101-2410-H-158-012) for sponsoring the compilation of LINDSEI-TW;
- Ministry of Science and Technology, Taiwan (grant number MOST 104- 2914-I-158 -003-A1) for sponsoring the presentation of this paper at JLC 2015;
- Dr Sheng Li for for his help in the matter of text extraction by using Bash scripts.

Thank you for your attention!

Lan-fen Huang

lanfen.huang@gmail.com

Shih Chien University, Kaohsiung, Taiwan