LUND UNIVERSITY

Stages of Processability and Levels of Proficiency in the Common European Framework of Reference for Languages. The Case of L3 French.

Granfeldt, Jonas; Ågren, Malin

[Link to publication](#)

Language Acquisition and Use in Multilingual Contexts

Theory and Practice

Anna Flyman Mattsson & Catrin Norrby (eds.)

Stages of processability and levels of proficiency in the Common European Framework of Reference for Languages

The case of L3 French

Jonas Granfeldt
Malin Ågren

Introduction

The work of Gisela Håkansson has for many years been an important source of inspiration for both of us. In particular her longtime work on L2 Swedish within PT has been decisive for us in previous attempts to discuss L3 French within this framework (Ågren 2008). The present paper is an extension of this work.

The aim of the present paper is to investigate empirically the relationship between *second language proficiency* (L2P) and *second language development* (L2D) in a corpus of written L3 French. Second Language Acquisition (SLA) has traditionally been concerned with describing and understanding L2D, most notably through the study of developmental sequences and stages. Language testers and language testing research are interested in capturing and measuring the broader concept of L2P at a given time. The question concerning a developmental relationship between L2P and L2D is not new and has been answered differently in the past. Within SLA, some researchers view L2D and L2P as separate theoretical constructs (Pienemann & Johnston 1987; Pienemann & Mackey 1992). R. Ellis (2008), for example, calls for attempts to match developmental levels and proficiency levels since he suspects that these two linguistic dimensions might in fact be a comparison of "apples and oranges" (Ellis 2008, note 7). The question we ask in this study is to what extent L2P and L2D develop in parallel in a group of L3 learners of French.

L2P can be defined as "a person's overall competence and ability to perform in L2" (Thomas 1994:330, footnote 1), to which we would like to add "at a given point in time" in order to underline the fluctuating and developmental aspects of L2P. Hulstijn (2011, 2012) has recently suggested a

subdivision of L2P into Basic Language Cognition (BLC) and Higher Language Cognition (HLC). The separation is motivated by the fact that more advanced aspects of L2P can be related to contextual (learning) factors which depend on the individuals' intellectual capacities and degree of formal schooling, rather than on purely linguistic skills. Higher levels of L2P cannot be reached, Hulstijn argues, without formal education. The distinction between BLC and HLC is thus needed in order to separate linguistic and intellectual skills as components of L2P (cf. Cummins, 1980, on a related division). BLC concerns the implicit knowledge and automated use of frequent and basic morphosyntactic constructions and lexical items.

L2P has been operationalized and measured in a variety of ways (see Hulstijn 2011) but there are essentially two broad approaches to measure L2P. On the one hand, the psychometric approach includes standardized tests (i.e. TOEFL, DELF), cloze tests, C-tests etc.[1] On the other hand, the holistic approach relies either on learners' self-rating or behavioral rating scales where expert raters assess L2 performance. In the holistic approach, criterion-referenced scales of language proficiency can be used and expert judges evaluate learner language by applying these scales, independently of target language. Currently, the most prestigious and well-known example is the *Common European Framework of Reference for Languages* (CEFR) which assesses communicative language proficiency. We return to the CEFR below.

Understanding and defining L2D is one of the core topics in SLA research. L2D can tentatively be defined as the progressive growth of one or more aspects of the interlanguage system (phonology, morphology, syntax, etc.). L2D could be viewed as a necessary but not sufficient subcomponent of L2P. It is often described via the definition of developmental sequences, which crucially are thought to be invariable and impermeable with respect to external factors such as learning situation, type of input, etc. An increasing number of different models and theories have been put forward to describe and account for L2D in different learners (Towell & Hawkins 1994; Pienemann 1998, 2005; Sharwood-Smith & Truscott 2005). Models of L2D tend to be language independent and some have also been empirically tested on cross-linguistic data. One well-known and established model of L2D is *Processability Theory* (PT) (Pienemann, 1998; Pienemann & Håkansson 1999) which over the years has been tested on a variety of typologically different languages. We return to PT below.

---

[1] TOEFL is also known as the Test of English as a Foreign Language. It is designed and administrated by the Educational Testing Service. The DELF test is the French equivalent. DELF is the acronym of Diplôme d'études en langue française. The test is designed and administrated by *Centre international d'études pédagogiques*. A C-test is a specific type of cloze-test used in language testing (Grotjahn, 2010).

The specific aim of the present study is to investigate empirically the possible relationship between CEFR, a model of communicative L2P, and PT, a model of L2D. The rationale is that the relationship between L2P and L2D is debated and needs further attention. With Hulstijn's (2011, 2012) division between BLC and HLC in mind, it seems important to understand at what point in the learners' trajectory L2D and L2P might be more or less associated. A possible hypothesis is that L2P in the BLC range would be closer associated with L2D since both are defined as constructs reflecting implicit knowledge of language and automated use. Moreover, both are independent from contextual and individual factors, such as the degree of formal schooling. Hulstijn criticises the CEFR for not being clear about the relationship between L2P and L2D and says that:

> Any association between CEFR levels of L2P [*L2 Proficiency*] and L2 development as studied in the second language acquisition (SLA) literature would be completely misplaced [...], unless empirical studies show evidence in its support. (Hulstijn 2011:241)

The remainder of this paper will be dedicated to a small-scale empirical study which aims to address this challenge.

The Common European Framework of Reference for Languages (CEFR)

The CEFR provides a language-independent description of communicative proficiency at six levels. The levels (A1, A2, B1, B2, C1 and C2) are organised according to three broad proficiency bands: Basic User (A), Independent User (B) and Advanced User (C). The CEFR is action-oriented; language learners are viewed as language users and as social agents who accomplish communicative activities. These activities all involve language in a broad sense and the CEFR describes *what* a learner can do with respect to a specific task at a certain level of communicative proficiency and *how well* s/he can do it.

Table 1. From *Overall written interaction* (Council of Europe, 2001, chapter 4).

CEFR Level	Descriptor
B1	Can convey information and ideas on abstract as well as concrete topics, check information and ask about or explain problems with reasonable precision.
A2	Can write short, simple formulaic notes relating to matters in areas of immediate need.

The CEFR encompasses four categories of language activities of this type: reception, production, interaction and mediation.

The scales of the CEFR define linguistic, pragmatic and socio-linguistic competences needed to carry out the activities. In particular chapter 5 presents communicative language competences (cf. Canale & Swain 1980). The way scales are presented in the CEFR could lead one to believe that functional and competence-based scales should be interpreted together. Hence, learners at, say, level A2 with respect to *Overall written interaction* (Council of Europe, 2001:83) should simultaneously be at the same level A2 with respect to competence-based scales in chapter 5, like for example *Vocabulary control* (Council of Europe 2001:112). In his paper, Hulstijn (2007:664) discusses such a "parallel" reading of CEFR functional and competence-based scales from a SLA perspective. At least three types of L2 language users can be identified: a) learners who can do few language tasks but with high linguistic quality, b) learners who can do many language tasks but with low linguistic quality and c) learners whose range of language tasks is parallel to their linguistic ability. Learner types a and b display "uneven profiles" but as a result of the way CEFR presents the scales, only the third learner type is included.[2]

Processability theory

Processability Theory (Pienemann 1998, 2005) is a psycholinguistic theory of SLA which explains developmental sequences in L2 acquisition in terms of language processing. According to this approach learners develop skills needed to process the target language grammar in a highly systematic way. Importantly, grammatical structures can only be produced in the L2 if the necessary processing procedures are available.

The processing hierarchy proposed in Pienemann's original version of PT, illustrated in Table 2, identifies five stages of development based on different levels of information exchanged between constituents (i.e. feature unification). The main idea is that the activation sequence of processing procedures used in *language production* (from 1 to 5 below) is also valid for *language acquisition*, which follows the same implicational order. Starting from stage 1, all subsequent stages of development mirror increasing demands of processing capacity involved in the morphosyntactic operations.

---

[2] On a single occasion the CEFR recognizes the existence of "uneven profiles" (Council of Europe 2001: 17), i.e. learners who are at different levels of proficiency in different activities, but nothing is said about the frequency or specificities of such learners.

Table 2. Stages of development and processing procedures according to PT (adapted from Pienemann, 1998).

PT stage	Processing Procedure	Information exchange	Example of morphol. outcome in French
5	Subordinate clause procedure	Main– sub. clause	Subjunctive in sub.clause
4	S(entence)procedure	Inter-phrasal	Subject-verb agreement
3	Phrasal procedure	Phrasal	NP agreement
2	Category procedure	No exchange	Lexical morphemes
1	Word or lemma access	Words, chunks	-

According to PT, the notion of storage of grammatical information in memory is crucial to the acquisition process. The further away the source and the target of feature unification, the longer grammatical information needs to be stored in memory, and the later the morphosyntactic structure will be acquired. Therefore, inter-phrasal agreement (stage 4) takes longer time to acquire than phrasal agreement (stage 3). The former asks for exchange of grammatical information over phrasal boundaries (between NP and VP), an operation that involves higher processing procedures, whereas the latter calls for more local information exchange within NP. As illustrated in previous cross-linguistic investigations of PT (see Pienemann 2005) the processing hierarchy affects morphology and syntax differently in various languages due to their different morphosyntactic rule systems. In this study, we focus on how this model applies to the development of morphosyntax in L3 French (Ågren 2008).

Research questions

We have two research questions for the present study:
- RQ1: To what extent is L2P, as measured by the CEFR, and L2D as defined by PT related in a corpus of written L3 French?
- RQ2: How frequent is the presence of uneven profiles in the data?

Data and method

The learners

All participants were pupils at different schools in the city of Lund in southern Sweden recruited through personal contacts with teachers at the schools. Written data were collected from 38 L3 learners of French. 36 of the 38 learners (95%) reported that Swedish was their mother tongue and the language used in the home. The learners were at two different levels: 22 in year 9 (15–16 years old, 59% female) and 16 in their final year of upper-

secondary school (18–19 years old, 69% female). All 38 learners successfully completed both tasks and were rated as being at least at the A1 level.

The tasks

A website was designed to be used for data collection. The students wrote their texts directly on the web page, without any kind of support. A total of 76 L3 French texts were collected (38 x 2) using this procedure. All learners were asked to complete two written communicative tasks. The tasks were adapted from two of the five tasks used by Alanen et al. (2010) in a study of young and adult learners' L2 English and L2 Finnish linked to the CEFR. Task 1 instructed students to write an email message to their French teacher explaining why they had been absent from school and asking for some information on an upcoming French test. The instructions to the older learners in upper-secondary school were similar but slightly more elaborated than those used in year 9 (for further details, see Granfeldt et al. in press). In task 2, students were asked to write a narrative about something nice, funny or special that they had experienced. They were instructed to explain what happened to them and why the event was particularly exciting or memorable. Participants were allowed 40 minutes to complete both tasks. All participants were able to complete the tasks within this time frame.

The CEFR raters

Two experienced CEFR raters were asked to read the texts and provide them with a CEFR score ranging from A1 to C2. As a basis for their assessment, the raters were given a CEFR scale that had been compiled from several of the CEFR scales. The compiled scale consisted of "can-do statements", and accuracy was never mentioned. The raters assessed the texts independently of each other and were asked to do the following: a) rate each text using the CEFR scale (A1, A2, B1, B2, C1 and C2), b) indicate the degree of certainty of each rating on a 4-point scale ranging from "completely certain" to "completely uncertain", and c) to provide an alternative CEFR rating if, and only if, the indicated degree of certainty was low (i.e. "uncertain" or "completely uncertain"). No texts were included in the final analysis where both raters were "completely uncertain" or "uncertain".

The degree of inter-rater agreement between the two raters was measured by Kronbach's alpha. Alpha was measured to .804 which according to DeVellis (1991:85) corresponds to a "very good reliability".

PT-analysis

One of the authors with previous experience of PT read the learner texts and analyzed them according to the PT framework. Since the tasks were adapted to match CEFR criteria, this meant that data density for some structures was low. Therefore, each analysis was evaluated on a 4-point scale according to the degree of certainty of the analysis. Out of the 76 CEFR rated texts, 61

texts contained a sufficient amount and varied set of structures in order to ensure a reasonable PT analysis with a certainty score of 3 or 4. These 61 texts were kept for the final comparison with the CEFR ratings.

Within the PT framework the *emergence criterion* is applied. Emergence refers to the first systematic and productive use of a certain structure, which indicates that the learner, in principle, can carry out a specific grammatical operation. In our PT analysis, the researcher studied the systematic use of certain structures based on three sources of available evidence in each learner text: 1) minimal pairs, 2) creative constructions (i.e. overuse of a grammatical rule), and 3) a certain amount of lexical variation in the use of a particular structure (e.g. the same morpheme used with a range of different lexemes). In the PT analysis, at least two minimal pairs *or* two creative constructions *or* three varied lexical items within a morphological pattern were required in order to identify a specific developmental stage (cf. Ågren 2008; Pallotti 2007; Pienemann 1998).

Results

In Figure 1 below, CEFR ratings of the 61 texts are plotted against the analysed PT-stage. Each learner text is represented by a circle in the figure. Since inter-rater agreement was estimated to be sufficiently high (see above), a single CEFR score for each text was computed by calculating a mean score from the two CEFR ratings. We observe that, overall, there is a linear correlation between the CEFR level and the PT stage.
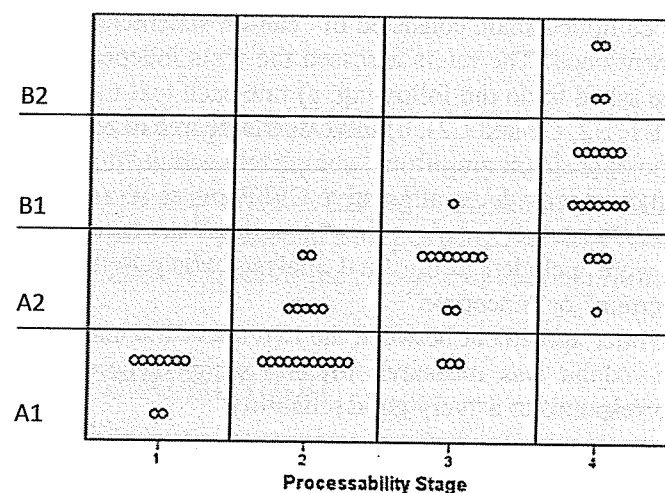

Figure 1. Correlation of CEFR rating and PT developmental stage.

The CEFR ratings cover four levels and range from A1 to B2. Texts that are plotted between two CEFR-levels have been rated differently by the two

raters. For example, a text which is plotted between A1 and A2 has been rated A1 by one rater and A2 by the other rater.

The results with respect to PT also cover four stages and range from PT stage 1 to stage 4. We observe that there is an increasing amount of variation in the relationship between the CEFR level and the analysed PT stage. According to the results in this small-scale pilot study, PT stages 1 and 2 are associated with CEFR levels A1 and A2. Interestingly, the dispersion increases at more advanced stages and PT stage 4 contain texts that have been rated from A2 (1 text) up to above B2 (2 texts). We also observe that, with a single exception, no text rated below B1 was analysed as PT stage 4.

To answer our first research question, we statistically investigated the strength of the association between the rated CEFR score (the mean) and the analysed PT stage using a Spearman rank order correlation analysis. The results indicate a very strong association between the average rated CEFR level and the analysed PT stage, $(rs[62] = .86, p<0.001)$.

In the remainder of this section we will address the second research question, namely the presence and frequency of uneven profiles.

For the purpose of this study we define a balanced profile as a text which follows the main linear trend expressed in Figure 1. These are texts where communicative proficiency (L2P), measured by CEFR, and morphosyntactic development (L2D), measured by PT, go hand in hand (cf. "flat profile", Council of Europe, 2001:43). We observe that balanced profiles seem to be dominant at lower levels and stages, which is not very surprising due to the limited command of the L2/L3 at these stages. Example (1) below illustrates a balanced profile at a more advanced level, where utterances are linked into a clear and coherent narrative (CEFR, level B2) and where the morpho-syntactic procedures at PT stage 4 have emerged (see phrasal agreement, stage 3, in solid underlining; interphrasal agreement, stage 4, in dashed underlining).

(1)   JL text 2: CEFR level B2 (both raters), PT stage 4

Il y a quelques semaines, je suis allée visiter une amie qui habite à Göteborg. En même temps, il y avait aussi ma meilleure amie de Lund qui allait visiter son amie à Göteborg. Nous, toutes les deux, sont donc allées en train pour Göteborg. En sortant du train à la gare de Göteborg, j'ai vu mon amie et on s'embrassées. Ma meilleure amie, elle a aussi rencontré son amie. Soudain, ma copine regarde la copine de ma meilleure amie et elle recoît un regard étonné. En fait, la copine de ma meilleure amie et ma copine de Göteborg étaient camarades de classe. Quand on a toutes compris cette coincidence sensationelle on s'est mises à rire...

However, it is not always the case that high/low communicative proficiency correlates with high/low morphosyntactic development. Texts showing un-even profiles appear as outliers in Figure 1. For instance, as exemplified in

(2), the learner ES's text was judged by both CEFR raters as belonging to level B1, where the learner is able to write accounts of experiences, describing feelings and reactions in simple connected text and link series of discrete elements into a connected, linear sequence of points. Indeed, ES is telling a rather straightforward narrative, where events from the past and the present are linked by the use of different tenses (imperfect and present tense), adverbials (*ne...jamais, la première fois, chaque fois...*) and pronominal reference (*la chanson... je l'ecoutais*). In Figure 1, CEFR B1 and B1+ and PT Stage 4 are closely associated. However, the PT analysis of ES' texts indicates stage 3, because the use of subject-verb agreement was not consistent enough to indicate availability of processing mechanisms at PT stage 4 (phrasal agreement in <u>solid underlining</u> has emerged, interphrasal agreement in <u>dashed underlining is questionable</u>). Thus, the communicative skills expressed in (2) seem to be more advanced than what could perhaps be expected given the stage of development.

(2) ES text 2: CEFR level B1 (both raters), PT stage 3

Je ne oblie jamais le premiere fois que je ecoutais à Edith Piaf. J'avait 12 ans et la musique en français etait très beau et je commencait de pleurer. C'est <u>la première fois</u> que je comprennait que je veux lire la français à l'école. La chanson etais "Paris" et c'est encore <u>ma chanson favorie</u> en français. Chaque fois j'ecoute la chanson je pense que la première fois que je l'ecoutais. C'est <u>une memoire très forte</u>...

Other texts, as exemplified in (3), show the opposite pattern where the level of morphosyntactic development was analysed at PT stage 4 (consistent <u>subject-verb agreement</u>) whereas the communicative proficiency was rated at a somewhat lower level (A2/B1) than could be expected (cf. the trend in Figure 1).

(3) TA text 2, CEFR level B1/A2, PT stage 4

<u>Je vais me souvenir</u> toujours de ce jour lorsque <u>moi et mes soeurs étions</u> avec vos amis à notre maison au bord de la mer. Là, il y a beaucoup de vaches. Un jour <u>nous avez oublié</u> le collier de notre chien dans un endroit où il y a des vaches. Quand <u>nous sommes allées</u> le chercher il y a telle de vaches dans cet endroit. <u>Nous pensions</u> que <u>cela était</u> très intéressant. <u>Ma petite soeur voulait monter</u> qu'<u>elle n'avait rien</u> peur et <u>elle était allé</u> proche d'un vache. Alors <u>la mère de cette vache était</u> furieux et <u>commençait</u> à courir après nous. <u>Nous avons couru</u> en pleurant. <u>Je n'ai jamais couru</u> si vite... Donc <u>tout allait</u> bien et maintenant c'est plutôt un bon souvenir.

Even though uneven profiles are present in the text sample examined in this study, we conclude that the majority of texts show rather homogeneous profiles, where L2P and L2D point in the same direction.

Summary

The present small-scale study aimed at investigating a possible relationship between second language proficiency (L2P) and second language development (L2D) in a corpus of L3 written French. The learners were Swedish secondary students. The learner texts were assessed according to the CEFR by two experienced CEFR raters, measuring communicative L2P. The same texts were analysed according to Processability Theory, which is a theory of L2D. We found a strong overall correlation between the CEFR ratings and the PT analysis ($rs[62] = .86$, $p<0.001$). We want to underline, however, that the observed correlation cannot in any way be taken as evidence that the underlying constructs are related.

In addition, we observed that the existence of uneven profiles in the data, i.e. learners with stronger communicative proficiency than morphosyntactic development or vice versa, typically becomes more frequent at more advanced stages. Up to CEFR B1 and PT stage 3, learners' communicative proficiency and morphosyntactic development seem to develop more or less at the same rate. This preliminary finding is potentially interesting in the view of L2P as divided between BLC and HLC (Hulstijn 2012). If BLC reflects implicit knowledge and automated language use, it actually comes a bit closer to the definition of L2D used in PT, which, in turn, could explain the better fit between CEFR and PT at lower levels and stages, as indicated in Figure 1. Future research will have to investigate this association in more detail.

References

Ågren, M. 2008. À la recherche de la morphologie silencieuse: sur le développement du pluriel en français L2 écrit. *Études Romanes de Lund*, 84. Lund University.

Alanen, R., Huhta, A. & Tarnanen, M. 2010. Designing and assessing L2 writing tasks across CEFR proficiency levels. In I. Bartning, M. Martin & I. Vedder (eds.), *Communicative proficiency and linguistic development: intersections between SLA and language testing research*. Eurosla Monographs: European Second Language Association. 21–56.

Canale, M. & Swain, M. 1980. Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1. 1–47.

Council of Europe, 2001. *Common European Framework of Reference for Languages: language, teaching, assessment*. Cambridge: Cambridge University Press.

Cummins, J. 1980. The construct of language proficiency in bilingual education. In J.E. Alatis (ed.), *Current issues in bilingual education (GURT 1980)*. Washington, DC: Georgetown University Press. 81–103.

DeVellis, R.F. 1991. *Scale Development*. Newbury Park: Sage Publications.

Ellis, R. 2008. Investigating grammatical difficulty in second language learning: Implications for second language acquisition research and language testing. *International Journal of Applied Linguistics* 18 (1). 4–22.

Granfeldt, J., Gyllstad, H. & Källkvist, M., in press. Linguistic correlates to comm-

unicative proficiency levels of the CEFR: The case of syntactic complexity in L2 English and L3 French. In C. Rosén, P. Simfors & A.K. Sundberg (eds.), *Rapport från ASLA:s vårsymposium*, Linköping, 2012.

Grotjahn, R. 2010. *Der C-Test: Beiträge aus der aktuellen Forschung*. Frankfurt am Main: Peter Lang.

Hulstijn, J. 2007. The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, 91. 662–667.

Hulstijn, J. 2011. Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second language assessment. *Language Assessment Quarterly*, 8. 229–249.

Hulstijn, J. 2012. The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, 15 (2). 422–433.

Pallotti, G. 2007. An operational definition of the emergence criterion. *Applied Linguistics*, 28 (3). 361–382.

Pienemann, M. 1998. *Language processing and second language development: Processability Theory*. Amsterdam: John Benjamins Publishing Company.

Pienemann, M. (ed.), 2005. *Cross-linguistic Aspects of Processability Theory*. Amsterdam: John Benjamins Publishing Company.

Pienemann, M. & Johnston, M. 1987. Factors influencing the development of language proficiency. In D. Nunan (ed.), *Applying Second Language Research*, Adelaide: NCRC. 45–142.

Pienemann, M. & Mackey, A. 1992. An empirical study of children's ESL development and Rapid Profile. In P. McKay (ed.), *ESL development. Language and literacy in schools*. Commonwealth of Australia and national languages and literacy institute of Australia. 115–259.

Pienemann, M. & Håkansson, G. 1999. A unified account toward the development of Swedish as L2. *Studies in Second Language Acquisition*, 21 (3). 383–420.

Sharwood-Smith, M. & Truscott, J. 2005. Stages or continua in second language acquisition: A MOGUL solution. *Applied Linguistics*, 26 (2). 219–240.

Thomas, M. 1994. Assessment of L2 proficiency in second language acquisition research. *Language Learning*, 44 (2). 307–336.

Towell, R. & Hawkins, R. 1994. *Approaches to second language acquisition*. Clevedon: Multilingual Matters.

So you think gestures are compensatory?

Reflections based on child and adult learner data[3]

Marianne Gullberg

Introduction

When I told Gisela that I wanted to examine how adult second language (L2) speakers used gestures as communication strategies in my doctoral dissertation, I remember her swallowing hard, then smiling and cheerfully saying *How exciting!* I thought it fitting to honour her by revisiting the can of worms I opened then, and take a new look at the question of whether, and if so how, gestures can be said to be compensatory.

In a seminal paper from 1985 entitled *So you think gestures are non-verbal?* David McNeill challenged the then dominant view of gestures as a communicative frill of no consequence to our understanding of language and linguistic processing (McNeill 1985). The paper listed arguments for why gestures are in fact verbal (i.e. linguistic), highlighting their close relationship with spoken language. Some 30 years later, this position has become well established. Evidence continues to accumulate for the close connection between gesture and language in language development, break-down, in processing, etc. Although the link itself is no longer questioned, the exact nature of the relationship and the reasons for why we gesture remain illusive.

It is a common lay assumption that speakers in expressive trouble use hand and foot solutions to resolve them. Gestures – particularly represent-ational or referential gestures which convey meaning about a referent (e.g. size, shape, etc.) – are seen as a compensatory tool to bridge the gap between communicative intention and available expressive means. This view is also common in research targeting language users who are "challenged" or "less competent". It is explicit in studies of adult second language acquisition and