The 'thinking' in thinking-for-speaking

Where is it?

Panos Athanasopoulos and Emanuel Bylund
Reading University / Stockholm University

According to the thinking-for-speaking (TFS) hypothesis, speakers of different languages think differently while in the process of mentally preparing content for speech. The aim of the present paper is to critically discuss the research carried out within the TFS paradigm, against the background of the basic tenets laid out by the proponents of this framework. We will show that despite substantial progress in the investigation of crosslinguistic differences in the organisation of information in discourse, the studies that actually examine the cognitive aspects of speech production are, to date, vanishingly few. This state of affairs creates a gap in our knowledge about the thought processes that co-occur with speech production during language use and acquisition. We will argue that in order to reach a more comprehensive picture of the cognitive processes and outcomes of speech production, methodologies additional to the analysis of information organisation must be used.

Keywords: Thinking-for-speaking, conceptualisation, linguistic relativity

1. Introduction

The idea that speakers of different languages think differently has been subject to scholarly debate and empirical inquiry in a range of different disciplines. Though this question was discussed already in the ancient schools of Greek and Sanskrit philosophy, it is commonly associated with the linguist Benjamin Lee Whorf, who formulated the so-called *linguistic relativity principle*. According to this principle, speakers of different languages are not equal observers of the world, but are pointed by their grammars towards different evaluations of reality (Whorf 1956: 241). Research into linguistic relativity thus seeks to establish whether crosslinguistic differences in the semantic partitioning of reality give rise to crosslinguistic differences in thought. A cornerstone in this line of inquiry is the definition of 'thought'

Language, Interaction and Acquisition 4:1 (2013), **91–100**. DOI 10.1075/lia.4.1.05ath
ISSN 1879–7865 / E-ISSN 1879–7873

as *non-verbal* behaviour, operationalised typically as a range of different cognitive processes, such as reasoning, classification, and categorical perception (Lucy 1992, 1997).

The linguistic relativity principle is, however, not the only account that provides a framework for research into the relationship between language and thought. In more recent years, Dan Slobin (1987, 1991, 1996, 2003) formulated the *thinking-for-speaking* (henceforth TFS) hypothesis as an alternative to the linguistic relativity principle. According to the TFS hypothesis, speakers of different languages think differently while in the process of mentally preparing content for speech. The crucial difference between LR and TFS is that the former focuses on effects of linguistic structure on non-verbal behaviour and conceptual representation, while the latter focuses on effects of linguistic structure on the cognitive processes involved in speech production, i.e. the conceptualisation stage in Levelt's (1989) terms. More specifically, Slobin suggests that the speaker attends to and verbalises those aspects of reality that are readily encodable in his/her language. The past decade has seen a steady increase in studies conducted within the TFS framework, and there is a growing number of investigations examining TFS in different types of bilingual speakers (for recent overviews and contributions, see Benazzo, Flecken & Soroli 2012; Cook & Bassetti 2011; Han & Cadierno 2011; Jarvis 2011; Pavlenko 2011).

The aim of the current paper is to critically discuss the research carried out within the TFS paradigm, against the background of the basic tenets laid out by the proponents of this framework. We will show that whereas there is a large number of crosslinguistic studies on the organisation of information in discourse, the studies that actually examine the cognitive aspects of speech production are, to date, vanishingly few. This bias creates a gap in our knowledge about the thought processes that co-occur with speech production, and reduces the 'thinking' component in TFS to a mere question of information structure. We will argue that in order to reach a more comprehensive picture of the cognitive processes and outcomes of speech production, methodologies additional to the analysis of information organisation must be used.

2. The thinking-for-speaking framework

The TFS framework postulates that the activity of thinking takes on a particular quality when used in the activity of speaking (Slobin 1996: 76). This cognitive process of thinking for the purpose of speaking constitutes, in other words, a "special form of thought that is mobilised for communication" (*ibid.*). In essence, when thinking for speaking, the speaker chooses those characteristics of events and

objects that, first, fit some conceptualisation of the event, and second, are readily encodable in his/her language (*ibid.*). In concrete terms, this means that, for example, when talking about a given event, a speaker of Turkish will have to check the properties of the information source underlying the predication, evidentiality being a grammatical category in Turkish, or that a speaker of Russian will have to check the temporal distribution of the event, since aspect is grammaticised in Russian.

The rationale underlying the TFS framework is that human beings are most of the time engaged in preparing, producing, or interpreting verbal messages, and therefore research into language and thought is incomplete without attention to the thought processes that relate to speech production. It is suggested that TFS effects can be documented in different time frames in relation to speech production (Slobin 2003: 183 *et passim*): the first frame concerns prelinguistic or nonlinguistic coding, which occurs when the speaker attends to those dimensions of reality that are relevant for subsequent linguistic coding. These effects are called 'anticipatory' effects. The second time frame is the speaking time proper, which is when the linguistically codable dimensions are transformed into speech (this time frame also applies to speech comprehension). The third frame is situated past speech production or comprehension, and relates to effects of the linguistic coding of reality on subsequent recall or recognition. These effects are called 'consequential' effects.

Slobin (1987) suggests that TFS can be divided into several subtopics, one of which concerns first language (L1) acquisition. For the past two decades, researchers have collected data on information structure in child and adult speakers in a number of different languages, with the intention to establish the emergence of language-specific patterns of TFS in the course of L1 acquisition (e.g., Berman & Slobin 1994; Strömqvist & Verhoeven 2004). Another TFS topic that is increasingly receiving attention relates to the phenomena of additional language learning. Here, the question at issue concerns whether or to what extent an additional language learner can attain the TFS patterns of the target language (e.g., Benazzo *et al.* 2012; Verhoeven & Strömqvist 2001).

3. The current state of TFS: Too much emphasis on the speaking, little attention to the thinking

What kinds of questions might be asked in a research agenda linking language and thought in native speakers and in bilinguals/L2 users? One aim would be to investigate how people think about and construct conceptual representations of objects, events, colours, spatial relations, time and the like as a function of their

native/second language. A second aim would be to address the thought processes taking place during speech production, probing the way speakers develop strategies of encoding and organising information during online speech production (or comprehension). The first aim is in fact at the heart of modern enquiries into the linguistic relativity hypothesis. As such, to address this aim, studies typically employ measures of nonverbal performance (e.g. categorisation, similarity judgements, recognition memory, scene comprehension). The second aim is closer to Slobin's TFS. According to Slobin (2003), to fully address this aim, some sort of methodology that would elucidate the temporal dynamics of developing strategies during speaking should be employed. Such methodologies might include speech onset time and comprehension reaction times, or some behaviour concurrent to speech production and/or comprehension (e.g. visual attention allocation, gesture, event-related brain potentials).

The first point we wish to raise in relation to the above aims is the importance of treating linguistic relativity and TFS as distinct (albeit conceptually related) hypotheses. While the cornerstone of both hypotheses is the fact of linguistic diversity in semantics and grammar, the methodologies employed to address each question are necessarily different (see previous paragraph and also Lucy 1997; Odlin 2011; Pavlenko 2011; Treffers-Daller 2012). From a theoretical viewpoint, it is important to distinguish these two hypotheses because they concern two different levels of mental representation. As Lucy (1997) and others have argued, phenomena related to selecting and structuring content for speech as captured by the thinking-for-speaking paradigm concern a linguistic level of representation distinct from a non-linguistic conceptual level that concerns cognitive representation of concepts, which may be language-derived to a lesser or greater extent depending on the conceptual category, the task, etc. (*cf.* Athanasopoulos & Bylund, 2013). Any kind of linguistic influence on that level of representation may be taken as evidence for linguistic relativity (Lucy 1992, 1997). There is ample empirical evidence to support the distinction between these two levels of processing. For example, studies show that cognitive representation of motion and thinking-for-speaking about motion are largely independent from each other. Across many different experiments, apparent cross-linguistic differences are found when talking about events, but not when cognitively processing motion scenes outside of overt verbalisation (Gennari, Sloman, Malt & Fitch 2002; Papafragou, Massey & Gleitman 2002) or memory-based cognitive processing (Papafragou & Selimis 2010; Trueswell & Papafragou 2010; Athanasopoulos & Bylund, 2013; Bylund, Athanasopoulos & Ooostendorp, in press). Clearly, the degree to which thinking-for-speaking depends on language-specific grammatical categories is orthogonal to the degree to which conceptual representations themselves are language-independent or not.

The second point we wish to put across in relation to the aims of this section, and most central for the purposes of the entire paper, is the observation that many studies purporting to investigate TFS *per se* do not in fact provide sufficient data on thought processes during speaking, but only describe linguistic diversity in the sense that they report typologically-constrained verbalisations produced by speakers of different languages. This, then, is taken as evidence of crosslinguistic differences in the thought processes concurrent to speech production. It is unclear, however, how much can be inferred about co-verbal thought processes from purely verbal data. In fact, this problem of inference resembles that which many early approaches to linguistic relativity faced, in which linguistic diversity was not only the premise, but also the final evidence of crosslinguistic differences in cognition. Such circular reasoning is parodied by Pinker (1994:61) as follows: "[Eskimos] speak differently so they must think differently. How do we know that they think differently? Just listen to how they speak!" Applying this logic to research into TFS, one could ask what evidence there is to assume that speakers of different languages think differently while describing a given scene because of the mere fact that they describe the scene differently. If the only evidence for this assumption is the crosslinguistic differences in information structure about the scene, then there is a clear component of circularity to this reasoning.

The literature abounds of studies examining semantic and syntactic aspects of information organisation from a crosslinguistic perspective (see contributions in, e.g., Benazzo *et al*. 2012; Berman & Slobin 1994; Han & Cadierno 2011; Pavlenko 2011; Strömqvist & Verhoeven 2004). The typical approach of studies of this kind is to compare verbal descriptions of stimuli amongst native speakers of different languages and/or language learners. The typical finding is that people tend to say what is typologically possible (in the case of obligatory linguistic constraints) and/or prototypical (in the case of non-obligatory but preferred linguistic features) in their native language, while L2 learners may show pervasive influence of the L1 in their L2 productions, but also in some cases influence of the L2 on the L1. While these studies may provide sophisticated linguistic analyses of information focus, semantic density, semantic locus, syntactic architecture and the such, they rely on speech elicitation methods that reveal little or nothing about the online psycholinguistic processes and dynamic conceptualisation strategies at play. Therefore, claims made about crosslinguistic differences in conceptual representation remain speculative. Production data alone reveal what people have said, and as such address only one aspect of Slobin's thinking-for-speaking hypothesis (speaking time proper). Such data tell us very little about how and why speakers choose to encode information in a particular way (Slobin's anticipatory stage), and little focus is paid to consequential effects on cognition of immediately preceding verbalisation (Slobin's third stage of thinking-for-speaking, see Gennari *et al*. 2002 for an

empirical approach). Because of the dynamic nature of speech production, the use of online methods of speech production performance (see next section for examples) would allow for a more complete picture of thinking-for-speaking processes.

4. Probing the thinking in TFS: Towards a new research agenda

In this section, we will argue for a methodological shift in focus from surface verbal production to co-verbal behaviours such as gestures, attention allocation during speech, and ERPs in semantic comprehension, because we believe it is in these co-verbal behaviours that the 'thinking' in thinking-for-speaking may be most usefully conceptualised, and empirically operationalised. The caveat of inferring information about conceptualisation processes from speech data alone is highlighted by the increasing amount of studies that have emerged in the past five years or so, showing that when dynamic measures of online behaviour are used, those measures can often reveal conceptualisation processes that are at variance with, or nuance further, surface production patterns.

For example, Flecken (2011) investigated effects of grammatical aspect on verbal encoding in early bilingual speakers of Dutch (an aspect language) and German (a non-aspect language). Participants had to describe a set of video clips focusing on the question 'What is happening?' Additionally, their eye movements were recorded in order to study their planning and organisation of content that was going to be expressed. The results showed that bilinguals displayed patterns of language use that were in-between the two monolingual norms. Specifically, these early bilinguals frequently used the progressive form in Dutch, but they also tended to combine progressive aspect with the mention of endpoints, a combination that is not at all frequent in monolingual Dutch speakers. The eye-tracking analysis revealed that the Dutch–German bilinguals, when speaking Dutch, have an underlying bilingual-specific event conceptualisation system that differs not only from that of Dutch monolinguals, but that also shows no sign of cross-linguistic influence from the monolingual German pattern.

Brown and Gullberg (2008) report dissociations in surface production and gesture behaviour in Japanese and English monolinguals and L2 users. When Japanese speakers mention manner of motion in speech, they are most likely to also gesture about manner. However, when English speakers mention manner in speech, they often gesture only about the path of motion. Japanese L2 English users were significantly more likely to talk about manner and gesture about path in their L1 Japanese than were monolingual Japanese speakers, thus displaying English-like conceptualization in their L1 productions. More strikingly, in a study of placement descriptions in the speech and gestures of Dutch and German

learners of L2 French, Gullberg (2011) found that both L2 groups produced target-like spoken L2 French with respect to the linguistic categories under scrutiny, suggesting restructuring of conceptualisation patterns as a function of L2 acquisition. However, when Gullberg (2011) examined the accompanying gesture behaviour of the learners, she found that both L2 groups showed evidence of L1 influence on conceptualisation, with many learners showing an in-between pattern, displaying both L1 and L2 gesture patterns within the same utterance.

To provide an example from comprehension processes, in a study looking at the effects of grammatical gender while having to make decisions about the semantic relationships between objects, Boutonnet, Athanasopoulos and Thierry (2012) tested Spanish-English bilinguals and native speakers of English in a semantic categorisation task. Participants were presented with triplets of pictures in an all-in-English context while measuring event-related brain potentials (ERPs). Participants were asked to press one button when the third picture of a triplet belonged to the same semantic category as the first two, and another button when it belonged to a different category. Unbeknownst to them, in half of the trials, the gender of the third picture name in Spanish had the same gender as that of the first two, and the opposite gender in the other half. Behavioural results showed no measurable effect of gender consistency in either of the participant groups. That is, Spanish-English bilinguals paid no more attention to gender incongruences in the triplets of stimuli than English monolinguals did. In contrast, ERPs revealed not only the expected semantic priming effect in both groups, but also a different brainwave pattern, modulated by gender inconsistency in Spanish-English bilinguals, exclusively. These results show that these bilinguals spontaneously and unconsciously access grammatical gender in their L1, even though on the surface they perform like native speakers of their L2.

In a study combining production patterns, attention allocation, and comprehension of scenes from memory, von Stutterheim, Andermann, Carroll, Flecken, and Schmiedtova (2012) showed that speakers of [+aspect] languages (e.g. Arabic, Spanish, English) who do not typically mention the endpoint of an action (Von Stutterheim & Nüse, 2003), also tend to look less at endpoints, and do not store information on endpoints to the same extent as speakers of [−aspect] languages (e.g. Czech, Dutch, German) do. The finding that speakers of [+aspect] languages mention more endpoints than speakers of [−aspect] languages shows an effect of linguistic typology on what speakers choose to mention. What shed light on the actual process of speech planning was the eye-tracking analysis: since the endpoint will not be relevant in planning what to say, speakers of [+ aspect] languages directed less visual attention to the endpoint-object in the stimulus than speakers of [−aspect] languages. Eye-tracking analyses also revealed that the effect of linguistic typology on conceptualisation is not an all-or-nothing phenomenon. Even

though speakers of [+ aspect] languages mentioned fewer endpoints than speakers of [–aspect] languages (as predicted by linguistic typology), visual attention in speakers of [+ aspect] languages was directed to the potential endpoint to some extent. The crucial difference between populations seems to be that [+ aspect] speakers look at the endpoints at a later point, compared to [–aspect] speakers (cf. von Stutterheim & Carroll 2006). In other words, it is necessary to also take into account the time course of visual attention as an additional window on conceptual processing. While [–aspect] speakers may direct attention to the endpoint region before speech onset, [+aspect] speakers may do so in the course of utterance articulation (von Stutterheim *et al.* 2012).

5. Conclusions

Linguistic diversity is the cornerstone of both the linguistic relativity hypothesis and the thinking-for-speaking hypothesis. However, linguistic diversity can never, in and of itself, be taken at face value as evidence for diversity in thinking (Lucy 1997) or in thinking-for-speaking (Slobin 2003). Purely linguistic analyses are a fundamental starting point for TFS, but at the same time a tad rudimentary as an only tool for ferreting out the cognitive processes involved in speech production. We have argued that the vast majority of studies look at the organisation of information structure in discourse when addressing TFS. But information structure is the end result of a complex process. Studies tend to overlook this process, or make assumptions about this process by looking at the end result, rather than at the process itself. Looking at speech data alone can tell us what people choose to say, but does not reveal how they come to say it. In other words, they leave the big question of the cognitive processes involved in preparing content for speech unanswered.

Acknowledgements

The work with this paper was supported by the Swedish Research Council (grant nr 421-2010-2104).

References

Athanasopoulos, P. & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science,* 37, 286–309.

Benazzo, S., Flecken, M. & Soroli, E. (eds.) (2012). *Typological perspectives on second language acquisition: 'Thinking for Speaking' in L2*. Special issue of *Language, Interaction, and Acquisition*, 3.

Berman, R. & Slobin, D. (eds.) (1994). *Relating events in narrative*. Hillsdale, NJ: Lawrence Erlbaum.

Boutonnet, B., Athanasopoulos, P. & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research* 1479, 72–79.

Brown, A. & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition* 30, 225–251.

Bylund, E., Athanasopoulos, P. & Ooostendorp, M. (in press). Motion event cognition and grammatical aspect: Evidence from Afrikaans. *Linguistics*.

Cook, V. & Bassetti, B. (eds.) (2011). *Language and bilingual cognition*. Hove, UK: Psychology Press.

Flecken, M. (2011). Event conceptualization by early Dutch-German bilinguals: Insights from linguistic and eye-tracking data. *Bilingualism: Language and Cognition* 14, 61–77.

Gennari, S., Sloman, S., Malt, B. & Fitch, W. (2002). Motion events in language and cognition. *Cognition* 83, 49–79.

Gullberg, M. (2011). Thinking, speaking, and gesturing about motion in more than one language. In A. Pavlenko (ed.), *Thinking and speaking in two languages* (143–169). Bristol: Multilingual Matters.

Han, Z. & Cadierno, T. (eds.) (2011). *Linguistic relativity in L2 acquisition: thinking for speaking.* Bristol: Multilingual Matters.

Jarvis, S. (ed.) (2011). *Crosslinguistic influence in bilinguals' concepts and conceptualizations*. Special issue of *Bilingualism: Language and Cognition*, 14.

Levelt, W. (1989). *Speaking: from intention to articulation*. Massachusetts: MIT Press.

Lucy, J.A. (1992). *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.

Lucy, J.A. (1997). Linguistic relativity. *Annual Review of Anthropology* 26, 291–312.

Odlin, T. (2011). On the interdependence of conceptual transfer and relativity studies. In Z. Han & T. Cadierno (eds.), *Linguistic relativity in L2 acquisition: Thinking for speaking* (183–194). Bristol: Multilingual Matters.

Papafragou, A., Massey C. & Gleitman, L. (2002). Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition* 84, 189–219.

Papafragou, A. & Selimis, S. (2010). Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes* 25, 224–260.

Pavlenko, A. (ed.) (2011). *Thinking and speaking in two languages*. Clevedon: Multilingual Matters.

Pinker, S. (1994). *The language instinct*. New York: William Morrow.Slobin, D. (1987). Thinking for speaking. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 13, 435–444.

Slobin, D. (1991). Learning to think for speaking: Native language, cognition, and rhetorical style. *Pragmatics* 1(1), 7–25.

Slobin, D. (1996). From "thought and language" to "thinking for speaking." In J. Gumperz & S. Levinson (eds.), *Rethinking linguistic relativity* (70–96). Cambridge: Cambridge University Press.

Slobin, D. (2003). Language and thought online: cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (eds.), *Language in mind: Advances in the study of language and thought* (157–192). Cambridge, MA: MIT Press.

Strömqvist, S. & Verhoeven, L. (eds.) (2004). *Relating events in narrative. Vol II. Typological and contextual perspectives*. Mahwah, NJ: Lawrence Erlbaum.

von Stutterheim, C., Andermann, C., Carroll, M., Flecken, M. & Schmiedtova, B. (2012). How grammaticized concepts shape event conceptualization in language production: Insights from linguistic analyses, eye tracking data, and memory performance. *Linguistics* 50, 833–867.

von Stutterheim, C. & Carroll, M. (2006). The impact of grammaticalised temporal categories on ultimate attainment in advanced L2-acquisition. In H. Byrnes, H.D. Weger-Gunthrap & K. Sprang (eds.), *Educating for advanced foreign language capacities: Constructs, curriculum, instruction, assessment* (40–53). Washington, DC: Georgetown University Press.

von Stutterheim, C. & Nüse, R. (2003). Processes of conceptualization in language production: Language-specific perspectives and event construal. *Linguistics* 41, 851–881.

Treffers-Daller, J. (2012). Thinking for speaking and linguistic relativity among bilinguals. *Language, Interaction, and Acquisition* 3, 288–300.

Trueswell, J.C. & Papafragou, A. (2010). Perceiving and remembering events cross linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language* 63, 64–82.

Verhoeven, L. & Strömqvist, S. (eds.) (2001). *Narrative development in a multilingual context*. Amsterdam: John Benjamins.

Whorf, B.L. (1956). *Language, thought, and reality. Selected writings of Benjamin Lee Whorf* (J.B. Carroll, ed.). Cambridge, MA: MIT Press.

Résumé

Selon l'hypothèse du "Penser-pour-parler" (Thinking-for-speaking ou TFS), les locuteurs de langues différentes n'organiseraient pas leur pensée de la même manière lors de la préparation mentale de leurs productions. L'objectif de cet article est de proposer une discussion critique des travaux effectués dans le paradigme du TFS, à partir des principes de base établis par ce cadre théorique. Nous montrons que, malgré les avancées importantes qu'a apportées l'étude des différences inter-langues dans l'organisation du discours, les travaux examinant réellement les aspects cognitifs de la production de la parole sont encore très rares. A partir de ce bilan, nous montrons l'insuffisance de nos connaissances concernant les processus cognitifs qui accompagnent la production pendant l'utilisation et l'acquisition du langage. Hormis l'analyse de l'organisation discursive, nous concluons que d'autres méthodologies sont nécessaires afin d'obtenir une vision plus globale des processus cognitifs présidant à la production du langage et des implications qui en émergent.