



ELSEVIER

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

# Journal of Memory and Language

journal homepage: [www.elsevier.com/locate/jml](http://www.elsevier.com/locate/jml)

## Editorial

### Same, different, or closely related: What is the relationship between language production and comprehension?



The historical tradition in psycholinguistics has largely been to study *either* language production or comprehension. Almost all of nineteenth century psycholinguistics, for instance, concerned the production of language, culminating in Wundt's two-volume *Die Sprache* of 1900. This also held for research in language acquisition which, largely based on diary data, almost exclusively concerned the child's production of speech until Eimas, Siqueland, Jusczyk, and Vigorito (1971) introduced the experimental study of speech perception in infants. During the 1970s psycholinguistics became almost exclusively comprehension research. Johnson-Laird opened his review of experimental psycholinguistics in the 1974 *Annual Review of Psychology* with the statement: "The fundamental problem of psycholinguistics is simple to formulate: what happens if we understand sentences?" Production research long remained an island of (quite sophisticated) speech error and hesitation analysis. But also when during the 1980s and 1990s the research agenda became more balanced between comprehension and production research, the two fields developed largely independently. Historically only a small set of studies have explicitly addressed the issue of relations between these two skills of the language user. These studies are briefly surveyed below.

Relations between production and comprehension were often invoked as causes of language change. Bredsdorff (1821, 1886) mentioned mishearing, misunderstanding, or misrecollection of sounds as potential causes of sound change in the speaker. Also indolence of the speaker would affect articulation and hence the sound patterns perceived by the listener, a view with which others, in particular Wundt (1900), concurred.

An important nineteenth century theory invoking the interaction of production and perception of speech is Wernicke's (1874) theory of self-monitoring in speech. During speech the sensory speech center is actively involved in controlling whether the spoken words match stored sound images. When the system breaks down, as in conduction aphasia, paraphasias will occur. Self-monitoring in speech

became a hot issue through Lee's (1950) discovery of the delayed auditory feedback (DAF) effect. When you hear your own speech delayed by some 150 ms, speech fluency dramatically breaks down. Based on these observations, Lee designed an engineering model of self-monitoring, which required feedback to take place within the syllable being spoken.

Almost simultaneously, Broadbent (1952) demonstrated that participants were unable to understand a new question while answering a previous question. Attention can focus on one or the other task, but not on both simultaneously. This insight led to Broadbent's famous filter model of selective attention. The issue kept returning in psycholinguistics. If we cannot simultaneously focus attention on listening to speech and producing it, these processes must involve access to the same processing resources. Which resources are involved here? Kempen (2000; see also Kempen, Olsthoorn, & Sprenger, 2012) argued that these are grammatical unification processes, which are shared in grammatical encoding and decoding. Where these processes happen to be identical, concurrent listening and speaking is indeed possible, as demonstrated in Marslen-Wilson's (1973) studies of close shadowing.

Scholars have often observed that relations between comprehension and production are affected by adaptation and learning. Levelt (2014) provides many examples. Wundt (1900, vol. I, p. 476), for instance, discusses the adaptations in L2 or dialect perception to native articulation: "each [perceived] speech sound is a complex in which one's own articulatory experience takes part". Conversely, it is an almost universal assumption in nineteenth century and in behaviorist twentieth century acquisition theory that the child exclusively learns by imitation. Perception leads production. Exceptions were noticed though. For instance, in the infant's babbling spontaneous production takes the lead, as authors such as Preyer (1882) and Tracy and Stimpfl (1894, 1909) argued. Wundt (1900) agrees, but dismisses claims by Darwin (1877), Taine (1876), and Preyer (1882) that the child can also invent

new words: “After all, children’s speech is a product of the child’s environment, in which the child itself is, essentially, only passively involved.” (Vol. I, p. 296). Behaviorist [Kantor \(1936\)](#) considered the listener as a “stimulus” to which the speaker “adjusts”. The mechanism underlying this adjustment however remains in the dark.

The discussion of relations between comprehending and producing language took a new turn following the “cognitive revolution” (see [Levelt, 1974/2008](#), Vol III for a detailed review). The premise that language users are endowed with linguistic “competence”, a generative grammar of their language, raised the question how that grammar was implemented in “performance”, in particular in speaking and listening. The initial idea was that linguistic rules match one-to-one to parsing operations in the listener and generative operations in the speaker. This “derivational theory of complexity” was at the basis of [Matthews’s \(1962\)](#) analysis-by-synthesis model, where the listener generates structures till a matching one is found, and of [Yngve’s \(1961\)](#) “depth hypothesis”, where a sentence’s complexity for the speaker/listener depends on the maximum number of phrase structure nodes that have to be put on hold when generating or parsing the sentence. This version of the derivational theory was soon replaced by the so-called “coding hypothesis”, first formulated by [Mehler \(1963\)](#). When you listen to or memorize a sentence you recode its surface form as deep structure plus transformations (such as “passive” or “negative”). Reproducing the sentence is going back from deep to surface structure. These are “psychologically real” linguistic structures, identical for comprehension and production. The complexity of mapping surface structure onto deep structure or inversely is a function of the number of transformations involved. A declarative sentence is easier to produce or understand than a passive one, which is harder than a passive negative one. Hence transformations are “psychologically real” in both comprehension and production. A flurry of experiments followed.

However, this transformational version of derivational complexity also had to be abandoned. [Fodor and Garrett \(1967\)](#) proposed to keep the coding hypothesis, i.e., the surface/deep structure mapping by the listener, but to forget about transformational complexity. The aim should rather be to find out which cues in the surface structure could be helpful in deriving deep structure. A relative pronoun can be such a cue because it indicates the existence of a subordinate clause in deep structure. (It is easier to parse *The man whom the dog bit died* than *The man the dog bit died*). The verb is especially informative because it signals the possible argument structures in deep structure. In other words, the listener uses a set of “perceptual strategies” for mapping surface onto deep structure. There is no obvious relation between this tool kit for comprehension and the speaker’s generative operations. This move introduced the era of listener-centered psycholinguistics, studying “what happens if we understand sentences”. It first led to relaxing the notion that listeners compute full surface structures, and later to giving up on listeners computing full deep structures. Rather, listeners compute the “gist” of a sentence, which is the conceptual state of affairs intended by the speaker.

Proposals for the architecture of the language user’s conceptual base emerged in the artificial intelligence world, for instance in work by [Schank \(1972\)](#) and [Winograd \(1972\)](#). These proposals often included sketches of parsers mapping input sentences onto conceptual structures, for which [Wood’s \(1970\)](#) augmented transition network (ATN) was an attractive formalism because it acknowledged the incrementality of sentence parsing and because the conditions on transitions in the network can be syntactic, semantic, or conceptual. However, no speaker models worth mentioning were proposed. The mostly tacit assumption was that speaking and listening involved a shared conceptual base. Interactions between the two modes of language use were hardly considered.

That was quite different in the domain of phonetics where, since the 1950s, efforts emerged to incorporate articulatory representations into phonetic models of speech perception. As [Lieberman \(1957\)](#) put it: “speech is perceived by reference to articulation” (p. 122). This had already been suggested by [Stetson \(1951\)](#). [Stevens \(1960\)](#) proposed a schematic “analysis-by-synthesis” model of speech perception in which the “spectral representation” auditorily derived from the speech signal is converted into an articulatory description, whose “active synthesis” yields a matching spectral representation. Lieberman’s life’s work was to develop and experimentally test his version of this analysis-by-synthesis approach, the “motor theory of speech perception”, fully reviewed in [Lieberman \(1996\)](#). Listeners map the acoustic signal onto the articulatory movements that could have produced the signal and from there onto the intended gestural pattern.

Speech perception affecting speech production has been a recurrent theme in the history of psycholinguistics. [Meringer and Mayer \(1895\)](#) already defined a category of “extraneous intrusions” in their analysis of slips of the tongue. A word just heard or seen could erroneously intrude into the utterance a speaker was planning, an observation time and again reconfirmed in modern speech error research. [Garrett \(1980\)](#) called them “environmental contaminants”. A similar phenomenon was observed by conversational analysts. [Schenkein \(1980\)](#) discussed a wealth of cases in which a word or phrase is repeated in immediately adjacent turns, such as *Does that sound wrong? – It sounds terrible*. A first experimental study of this phenomenon showed that it can even be induced for adverbs without obvious semantic, pragmatic, or conversational functions ([Levelt & Kelter, 1982](#)). The causation seems to be one of mere “mechanical” priming from perception to production, as in the case of environmental speech errors. But the mediating mechanism remained unexplained.

In an independent development, however, perception-to-production priming became a powerful tool in the study of lexical access and syntactic planning in production. [Glaser and Döngelhoff \(1984\)](#) introduced the picture–word interference paradigm, where named target pictures were accompanied by visual distractor words, which could be semantically related to the target word. [Schriefers, Meyer, and Levelt \(1990\)](#) extended the paradigm by using spoken distractor words, which could be phonologically or semantically related to the target word. Picture naming

latencies were systematically affected by such priming at different stimulus onset asynchronies.

Around the same time [Bock \(1986\)](#) introduced the syntactic priming paradigm, where the participant repeats a prime sentence presented by the experimenter, and then describes a picture of an event (e.g., a dog chasing a mailman). The syntactic form used in describing the picture tends to follow the syntax of the prime (such as passive or active). The same result is obtained when the participant only hears the prime sentence, as shown in later experiments. These interference and priming paradigms have yielded a solid body of evidence for speech perception affecting speech production. As [Levett \(2001\)](#) pointed out, the grand challenge for the research community is to account for these effects by unified theories of the speaker/listener.

Over the last couple of decades or so, unified models of speaking, listening, and their relationship have in fact begun to be developed. Most of these theories have focused on prediction involving the speech production system (in line with recent evidence that reading and listening heavily draw upon predictive processes), and in several frameworks, most notably those by [Dell and Chang \(2013\)](#), [Pickering and Garrod \(2013\)](#), [McCauley and Christiansen \(2011\)](#).

[Dell and Chang \(2013\)](#) suggest that language comprehension involves prediction using the production system. In their P-chain framework engaging prediction leads to prediction error. This error signal is assumed to drive learning during language development and also explains performance in mature language processing. Structural priming in adults on this account occurs because these error-based learning mechanisms stay on in proficient language users. [Pickering and Garrod \(2013\)](#) propose that language users use forward production models in a similar way that actors use forward action models (cf. [Wolpert, Doya, & Kawato, 2003](#)). According to this view speakers construct efference copies of their predicted productions. These efference copies are compared with the output of a production implementer. Listeners are assumed to use these forward production models and covertly imitate speakers to predict upcoming utterances. [McCauley and Christiansen \(2011\)](#) argue that language production and comprehension are a single system. Syntactic knowledge accumulates through abstraction over multi-word sequences, and words are chunked based on transitional probabilities as incoming utterances are processed. The distributional information of the chunks employed during production is used to predict upcoming language input during comprehension.

These broad models are useful as they provide unified frameworks in which to place empirical findings and theories concerning specific aspects of language production or comprehension. But it can hardly be said that with their arrival the question of how production and comprehension are related has been answered. What is needed now is further detailed empirical and theoretical work into the way language production and comprehension are related in different components of the linguistic system, say at the phonetic, phonological, morphological or semantic level. This is because the relationship between production and

comprehension is not only important in its own right – for understanding the overall architecture of the cognitive system serving language – but also because assumptions about the relationship between the speech production and comprehension system have far-reaching consequences for theories of each of the individual language skills. For instance, when representations or processes are postulated to be shared between production and comprehension, there need to be processes that configure these shared components in such a way that they support either production or comprehension (e.g., phonological encoding or decoding), as need may be. Conversely, when representations are postulated to be unique for production or for comprehension, processes need to be invoked that explain how these distinct production and comprehension representations are acquired. This is not trivial given that people learn to speak from listening to others. The papers in this Special Issue contribute to this enterprise – the detailed exploration of the relationships between production and comprehension – in a number of complementary ways. Rather than focusing on production or comprehension of language they all place the relationship between the two skills in the center of the stage.

Four papers concern language learning in adults, specifically the way comprehension affects production or vice versa. [Kittredge and Dell \(2016\)](#) report three experiments on the acquisition of novel phonotactic constraints (e.g., “/f/ appears only as a syllable onset, not as a coda”) in a speech production task. They examined whether experiencing the same or an opposing constraint (e.g., “/f/ appears only as a coda”) in another person’s speech affected the participants’ speech output (specifically the types of errors committed). They found that the nature of the constraint in the speech input only mattered when participants subvocally produced the other person’s utterances, or when they monitored the utterances for errors. There was no transfer of learning when participants listened to the speech input and monitored for the occurrence of certain phonemes. The authors propose that the representations involved in production and comprehension are separate, but that some comprehension tasks (e.g., monitoring for errors) activate the production system. When this happens, transfer occurs from comprehension to production.

Whereas [Kittredge and Dell](#) examined the effect of speech input on speech production, [Baese-Berk and Samuel \(2016\)](#) studied the effect of speaking on the processing of the speech input, specifically on learning a novel sound contrast. Somewhat counter-intuitively, they found that overt repetition of the input syllables and, to a lesser extent, producing unrelated syllables disrupted learning of the novel sound contrast compared to a listening-only condition. As the authors discuss, the requirement to produce speech in addition to listening may increase the general processing load, and, more interestingly, there may be competition between novel and existing representations of sound forms, which become activated during the production task.

[Van Assche, Duyck, and Gollan \(2016\)](#) report two repetition priming experiments, one conducted in the speakers’ L1 and one in their L2, examining cross-modal transfer,

from repeated picture naming to lexical decision for the same items or vice versa. The robust transfer effects point to the existence of shared, or closely linked lexical representations for production and comprehension.

Zamuner, Morin-Lessard, Strahm, and Page (2016) studied the acquisition of new labels for new concepts (pictures of nonce-animals). In the training phase of their experiments, participants either heard the labels while looking at pictures of the referents or heard the labels and also repeated them. On test trials, they saw two of the pictures, heard the name of one of them, and were asked to look at that target. Fixation patterns indicated a beneficial effect of producing the object names during training, compared to only hearing them. In line with the view put forward by Kittredge and Dell, the authors conclude that producing the labels during training leads to the creation of production-based representations with bidirectional links to perception-based representations of the same items. These additional representations and their links to the perception-based representations support the correct mapping between sounds and pictures.

Buz, Tanenhaus, and Jaeger (2016) report a web-based study demonstrating how speakers change their articulation depending on the context and the communicative success of earlier utterances. Speakers hyper-articulated object names when the display shared with the (simulated) interlocutor featured competitor objects with phonologically similar names, especially when the interlocutor's non-verbal behavior (the selection of an incorrect referent object) indicated that they had misunderstood phonetically similar utterances on earlier trials. In this study, the participants' speech was not affected by verbal input from another person but rather by non-verbal feedback. The results support the existence of highly sophisticated speech monitoring and self-correction processes that allow speakers to adapt their articulation in subtle and specific ways based on feedback received from a listener.

Whereas the first group of papers concerns online learning processes, the second group concerns mediated longer-term effects of production processes onto comprehension processes. As MacDonald (2013) has pointed out, the ease of comprehending different utterance structures depends to a substantial degree on the listeners' or readers' experience with similar utterances. How often readers and listeners encounter different structures depends on their prevalence in spoken and written texts, which in turn depends, to a degree, on speakers' and writers' preferences. Thus, structures that speakers prefer because they are easy to produce will occur relatively frequently in spoken and written texts, and will therefore be expected and processed more easily than structures that speakers tend to avoid. Support for this view comes from three papers included in the volume. They all used corpus analyses to document speaker preferences and reading or listening tasks to establish the ease of processing different utterance types.

In the first of these papers Hsiao and MacDonald (2016) report four studies on the production and comprehension of Mandarin relative clauses. Production preferences (object relative versus passive relative constructions for phrases referring to animate or inanimate event

participants) were assessed through a picture description experiment and a corpus study. Comprehension was assessed in a gated sentence completion and self-paced reading experiment. The results provide strong evidence that the processing costs in sentence reading depend on the reader's prior experience with different structures and families of sub-structures defined both in terms of semantic and syntactic characteristics. The availability of these different structures can, to a substantial degree, be traced back to constraints arising when speakers and writers formulate utterances.

The next two papers concern code-switching. Fricke, Kroll, and Dussias (2016) showed through corpus analyses that code-switches from English to Spanish were preceded by subtle phonetic changes, namely slower overall speech rate and more Spanish-like VOT, compared to matched speech without code-switches. They attributed these changes to the increased cognitive demands that arise when a code-switch is planned and to simultaneous anticipatory activation of the incoming language and the deactivation of the current language. In a visual world eye tracking study the authors showed that bilingual listeners presented with code-switched utterances, which did or did not feature the characteristic pre-switch slowing and VOT changes, used these cues, where available, in processing the utterances.

The study by Guzzardo Tamargo, Kroff, and Dussias (2016) focused on the preferred locations of switches from Spanish to English in perfect and participle constructions. In their corpus study the authors established, in line with earlier studies, that switching between the auxiliary and the verb form was quite common for participle constructions ("están enjoying") but not for perfect constructions (\*"han enjoyed"), where the auxiliary and following verb form are rarely separated. The authors link this difference in the preferred switch locations to differences in the time course of retrieving the auxiliaries during grammatical encoding. In an eye movement study they demonstrated that the speakers' structural preferences were paralleled in reading times for constructions featuring code switches in expected (i.e., speaker preferred) and unexpected locations. The readers were sensitive to the distributional patterns in the language, which derived from production constraints, and used this information during reading. Thus, this paper, as the papers by Hsia and MacDonald and by Fricke, Kroll and Dussias, demonstrates how parallels between production and comprehension preferences can arise, mediated through the input that speakers and writers provide for listeners and readers.

Using a related research strategy, Gahl and Strand (2016) determined the influence of lexical and perceptual properties of words on their spoken durations in a corpus of connected speech and assessed the impact of the same word properties on listeners' performance in a perceptual task, word recognition in noise. They found that a lexical variable, segment-based phonological neighborhood density, affected both word recognition performance and word durations, whereas a measure of perceptual similarity (Phi-density) only affected word recognition, but not spoken durations. Thus, word durations reflect word-level information and the ease of phonological encoding of

words, and are not affected by the perceptual similarity of words to their neighbors.

Two papers focus on the role of domain-general processes during speaking and listening. The first of these papers, by [Ünal and Papafragou \(2016\)](#), is a developmental study and concerns the acquisition of evidential morphology markers (i.e., morphemes specifying the source of knowledge, i.e., direct or indirect experience). In a series of experiments, the authors demonstrate that children produce the appropriate morphological forms before they can correctly interpret them in utterances they hear. They argue that this pattern is not due to inherent differences in morphological processing in speech production and comprehension but rather to the interplay of these processes with perspective-taking, which is harder during comprehension than during production. This paper illustrates that speaking and listening cannot be seen as ‘the same processes running in opposite directions’.

The second paper focusing on a domain-general process, by [Humphreys, Mirkovic, and Gennari \(2016\)](#), concerns the mechanisms involved in resolving similarity-based semantic competition. The authors provide empirical evidence that semantic similarity elicits competition during planning, results in comprehension difficulty in certain syntactic configurations, and also influences the choice of syntactic structure in language production. The authors argue that a shared semantic competition mechanism underlies both comprehension and production and thus go beyond the simple claim that similar mechanisms play ‘some’ role in both processes.

The final two papers describe computational approaches of the relationship between production and comprehension. [Roon and Gafos \(2016\)](#) present a dynamical model of phonological planning. They use evidence from the response-distractor task (in which participants learn pairs of visual cues and spoken syllables, and hear distractor syllables when preparing a required response) to model how response planning is influenced by perception. The authors use the principles of excitation and inhibition in an explicit computational account to show how perceived stimuli can influence the planning process. They also describe how these principles can account for a number of other experimental findings such as effects of within- and across-category variation, multiple (mis)articulations, and VOT modulation.

[Chater, McCauley, and Christiansen \(2016\)](#) propose that language is best viewed in terms of skills rather than abstract representations. Their model of a Chunk-Based Learner represents an integrated approach of language production and comprehension. In contrast to most authors in this special issue they argue that the same representations and processes are involved across all linguistic levels. Chater et al. suggest that there is no abstract linguistic knowledge (e.g., no mental lexicon containing phonological and semantic properties of words) that is independent of language processing operations. Language learning therefore does not involve learning abstract language structure but is acquiring perceptuo-motor skills.

In sum, most papers in the volume address the issue of how production and comprehension processes affect each other, either directly (in on-line tasks), or more indirectly,

with speaker preferences shaping the experience available to listeners. Each paper raises interesting new questions. In our view a particularly pressing issue is how people shape each other’s language in actual conversation. For instance, in the first two papers of the volume, Kittredge and Dell show that in laboratory experiments speakers pick up phonotactic constraints in another person’s speech; [Baese-Berk and Samuel \(2016\)](#) show that producing syllables interferes with sound learning. These are important results. In order to understand how language ‘actually’ or ‘in the wild’ works, it is crucial to establish the impact of different learning mechanisms in everyday situations, where speakers interact with other people in joint activities (rather than focusing on the language tasks per se) and where they are exposed to many different interlocutors. How powerful are, for instance, the error-based learning mechanisms? Do people predict (and therefore learn) in ALL conversations? To what extent are predictions based on the speakers’ own production rather than on their previous experience with other persons? The question of how production and comprehension are related is intimately linked to the question of mutual influence and alignment of people in dialogue.

However, dialogue, though clearly important, isn’t everything. Several authors of the present volume used written text corpora in their empirical work and highlight the importance of exposure to written texts for shaping language skills. An important and challenging question for further research is how much spoken and written input people receive, and what the impact of this information is on comprehension, prediction, and production. It is plausible that exposure to written texts affects reading skills. If production and comprehension processes and representation are tightly linked experience with written texts should also affect speaking skills. Perhaps this is the case, but if so, the influence may be rather limited. Intuition suggests that you need to read a lot of Dickens to begin to talk like him.

Another issue taken up in several papers in the current volume concerns the interaction between linguistic processes and domain general processes. As the studies by [Humphreys](#) and colleagues and by [Unal and Papafragou](#) demonstrate, domain general processes shape production and comprehension. [Humphreys](#) and colleagues argue that resolving competition plays a similar role in production and comprehension. One may ask which other cognitive control processes have similar pervasive effects, and which processes have selective effects on particular components of speech planning or comprehension. [Unal and Papafragou](#) illustrate how a comprehension deficit (relative to production) arises in children due to the interaction between morphological processing and the recruitment of perspective taking skills. It would be interesting to see whether similar patterns arise elsewhere, for instance in the use of deictic pronouns or personal pronouns, and more generally, to what extent asymmetries between production and comprehension skills are conditioned by the recruitment of domain-general processes or the use of linguistic knowledge. This is not only important for understanding these puzzling asymmetries, but, more importantly, also for understanding the way linguistic and general processing components co-operate in developing and adult speakers.

Readers who do not wish to read all papers of the volume (which of course, we strongly recommend) may wonder whether there is any consensus among the authors regarding the simple question whether speaking and listening share, or do not share, representations. Kittredge and Dell distinguish between positions that postulate separate, inseparable, and separable production and comprehension representations. The separate position, which implies no mutual influences between production and comprehension representations and processes, is not endorsed by any of the authors of the Special Issue. Chater and colleagues proposed a shared set of skills that support both speaking and listening. The remaining papers discussing the relationship (not all did) subscribed to the 'separable' view, according to which production and comprehension representations and processes are separate (distinguishable theoretically and empirically) but can influence each other under certain conditions. Thus, the authors of the volume appear to be in good agreement that speech production and comprehension engage skills and representations that are distinct but tightly linked.

### Author notes

The relationship between speech comprehension and production was discussed during a workshop held at the Max Planck Institute for Psycholinguistics in Nijmegen in March 2014. We thank all participants, especially Eve and Herb Clark, for their contributions.

### References

- Van Assche, E., Duyck, W., & Gollan, T. H. (2016). Linking recognition and production: Cross-modal transfer effects between picture naming and lexical decision during first and second language processing in bilinguals. *Journal of Memory and Language*, 89, 37–54.
- Baese-Berk, M. M., & Samuel, A. G. (2016). Listeners beware: Speech production may be bad for learning speech sounds. *Journal of Memory and Language*, 89, 23–36.
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387.
- Bredsdorff, D. J. H. (1821/1886). *Om Aarsagerne til Sprogenes Forandring*. Copenhagen: Gyldendalske.
- Broadbent, D. E. (1952). Speaking and listening simultaneously. *Journal of Experimental Psychology*, 43, 267–273.
- Buz, E., Tanenhaus, M. K., & Jaeger, T. F. (2016). Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory and Language*, 89, 68–86.
- Chater, N., McCauley, S. M., & Christiansen, M. H. (2016). Language as skill: Intertwining comprehension and production. *Journal of Memory and Language*, 89, 244–254.
- Darwin, Ch. (1877). A biographical sketch of an infant. *Mind. A Quarterly Review of Psychology and Philosophy*, 2, 285–294.
- Dell, G. S., & Chang, F. (2013). The P-chain: relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634), 20120394.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303–306.
- Fodor, J. A., & Garrett, M. (1967). Some syntactic determinants of sentence complexity. *Perception & Psychophysics*, 2, 289–296.
- Fricke, M., Kroll, J. F., & Dussias, P. E. (2016). Phonetic variation in bilingual speech: A lens for studying the production-comprehension link. *Journal of Memory and Language*, 89, 110–137.
- Gahl, S., & Strand, J. F. (2016). Many neighborhoods: Phonological and perceptual neighborhood density in lexical production and perception. *Journal of Memory and Language*, 89, 162–178.
- Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production. Speech and talk* (Vol. 1, pp. 177–220). London: Academic Press.
- Glaser, W. R., & Döngelhoff (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 640–654.
- Guzzardo Tamargo, R. E., Kroff, J. R. V., & Dussias, P. E. (2016). Examining the relationship between comprehension and production processes in code-switched language. *Journal of Memory and Language*, 89, 138–161.
- Hsiao, Y., & MacDonald, M. C. (2016). Production predicts comprehension: Animacy effects in Mandarin relative clause processing. *Journal of Memory and Language*, 89, 87–109.
- Humphreys, G. F., Mirkovic, J., & Gennari, S. P. (2016). Similarity-based competition in relative clause production and comprehension. *Journal of Memory and Language*, 89, 200–221.
- Johnson-Laird, P. N. (1974). Experimental psycholinguistics. *Annual Review of Psychology*, 25, 135–160.
- Kantor, R. (1936). *An objective psychology of grammar*. Bloomington: Principia Press.
- Kempen, G. (2000). Could grammatical encoding and grammatical decoding be subserved by the same processing module? *Behavioral and Brain Sciences*, 23, 38–39.
- Kempen, G., Olsthoorn, N., & Sprenger, S. (2012). Grammatical workspace sharing during language production and language comprehension: Evidence from grammatical multitasking. *Language and Cognitive Processes*, 27, 345–380.
- Kittredge, A., & Dell, G. S. (2016). Learning to speak by listening: Transfer of phonotactics from perception to production. *Journal of Memory and Language*, 89, 8–22.
- Lee, B. S. (1950). Effects of delayed speech feedback. *Journal of the Acoustical Society of America*, 22, 824–826.
- Levelt, W. J. M., *Formal grammars in linguistics and psycholinguistics*. Vol III, *Psycholinguistic applications*, 1974, The Hague: Mouton (re-edited, 2008, Amsterdam: Benjamins).
- Levelt, W. J. M. (2014). *A history of psycholinguistics: The pre-Chomskyan era* (2nd ed.) Oxford: Oxford University Press.
- Levelt, W. J. M. (2001). Relations between speech production and speech perception: Some behavioral and neurological observations. In E. Dupoux (Ed.), *Language, Brain and Cognitive Development: Essays in Honour of Jacques Mehler* (pp. 241–256). Cambridge, MA: MIT Press.
- Levelt, W. J. M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, 14, 78–106.
- Liberman, A. M. (1957). Some results of research on speech perception. *The Journal of the Acoustical Society of America*, 29, 117–122.
- Liberman, A. M. (1996). *Speech: A special code*. Cambridge, MA: MIT Press.
- MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, 4, 226.
- Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, 244(5417), 522–523.
- Matthews, G. H. (1962). *Analysis by synthesis of sentences of natural languages. Proceedings of the international congress on machine translation and applied language analysis*. London: H.M.S.O.
- McCauley, S., & Christiansen, M. H. (2011). Learning simple statistics for language comprehension and production: The CAPPUCCINO model. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 1619–1624). Austin, TX: Cognitive Science Society.
- Mehler, J. (1963). Some effects of grammatical transformation on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 2, 346–351.
- Meringer, M., & Mayer, C. (1895). *Versprechen und Verlesen. Eine psychologisch-linguistische Studie*. Stuttgart: Göschen'sche Verlagshandlung (Re-edition by A. Cutler & D. Fay (Eds.) (1978). Amsterdam: Benjamins).
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36(04), 329–347.
- Preyer, W. (1882). *Die Seele des Kindes. Beobachtungen über die geistige Entwicklung des Menschen in den ersten Lebensjahren*. Leipzig: Th. Grieben's Verlag.
- Roon, K., & Gafos, A. I. (2016). Perceiving while producing: Modeling the dynamics of phonological planning. *Journal of Memory and Language*, 89, 222–243.
- Schank, R. C. (1972). Conceptual dependency: A theory of natural language understanding. *Journal of Cognitive Psychology*, 3, 532–631.
- Schenkein, J. (1980). A taxonomy of repeating action sequences in natural conversation. In B. Butterworth (Ed.), *Language production. Speech and talk* (Vol. 1, pp. 21–48). New York/London: Academic Press.

- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86–102.
- Stetson, R. H. (1951). *Motor phonetics*. Amsterdam: North-Holland Publishing Company.
- Stevens, K. (1960). Toward a model for speech recognition. *The Journal of the Acoustical Society of America*, 32, 47–55.
- Taine, H. (1876). De l'acquisition du langage chez les enfants et dans l'espece humaine. *Revue Philosophique*, 1, 23–32.
- Tracy, F., & Stimpfl, J. (1894/1909). *The psychology of childhood*. Boston: Heath & Co.
- Ünal, E., & Papafragou, A. (2016). Production-comprehension asymmetries and the acquisition of evidential morphology. *Journal of Memory and Language*, 89, 179–199.
- Wernicke, C. (1874). *Der aphasische Symptomencomplex. Eine psychologische Studie auf anatomischer Basis*. Breslau: Max Cohn & Weigert.
- Winograd, T. (1972). Understanding natural language. *Cognitive Psychology*, 3, 1–191.
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society: Biological Sciences*, 358, 593–602.
- Woods, W. A. (1970). Transition network grammars for natural language analysis. *Communications of the ACM*, 13, 591–606.
- Wundt, W. (1900). *Die Sprache* (Vols. 2) Leipzig: Engelmann.
- Yngve, V. H. (1961). The depth hypothesis. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language*. Englewood Cliff, N.J.: Prentice-Hall.
- Zamuner, T. S., Morin-Lessard, E., Strahm, S., & Page, M. P. A. (2016). Spoken word recognition of novel words, either produced or only heard during learning. *Journal of Memory and Language*, 89, 55–67.

Antje S. Meyer

Max Planck Institute for Psycholinguistics, Nijmegen,  
The Netherlands  
Radboud University, Nijmegen, The Netherlands  
E-mail address: [Antje.Meyer@mpi.nl](mailto:Antje.Meyer@mpi.nl)

Falk Huettig

Max Planck Institute for Psycholinguistics, Nijmegen,  
The Netherlands

Willem J.M. Levelt

Max Planck Institute for Psycholinguistics, Nijmegen,  
The Netherlands  
Radboud University, Nijmegen, The Netherlands

Available online 22 March 2016